DISTRIBUTED NETWORK SYNCHRONIZATION: THE INTERNET AND ELECTRIC POWER GRIDS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Enrique Mallada

January 2014

DISTRIBUTED NETWORK SYNCHRONIZATION:

THE INTERNET AND ELECTRIC POWER GRIDS

Enrique Mallada, Ph.D.

Cornell University 2014

Synchronization is a fundamental requirement of most networked engineering applications. It enables the necessary coordination among agents required to implement several communication systems as well as network protocols. Despite the great recent advances in understanding synchronization, a complete synchronization theory is yet to be developed. This thesis presents a systematic study of synchronization on distributed systems that covers theoretical guarantees for synchronization, performance analysis and optimization, as well as design and implementation of algorithms.

We first present several theoretical results that deepen the understanding of how coupling, delay and topology affect the behavior of a system of coupled oscillators. We obtain a sufficient condition that can be used to check limit cycle stability, and use it to characterize a family of coupling functions guaranteeing convergence to in-phase synchronization (phase consensus). The effect of heterogeneous delay is then investigated by developing a new framework that unveils the dependence of the orbit's stability on the delay distribution. Finally, we consider the effect of frequency heterogeneity. While coupled oscillators with heterogeneous frequency cannot achieve phase consensus, we show that a second order version of the system can achieve synchronization for arbitrary natural frequencies and we relate the limiting frequency of the system to the harmonic mean of the natural frequencies.

Based on the insight provided by our theoretical results, we then focus on more practical aspects of synchronization in two particular areas: information networks and power networks. Within information networks, we examine the synchronization of computer clocks connected via a data network and propose a discrete algorithm to synchronize them. Unlike current solutions, which either estimate and compensate the frequency difference (skew) among clocks or introduce offset corrections that can generate jitter and possibly even backward jumps, this algorithm achieves synchronization without any of these problems. We present a detailed convergence analysis together with a characterization of the parameter values that guarantee convergence. We then study and optimize the effect of noisy measurements and clock wander on the system performance using a parameter dependent $\mathcal{H}_2$ norm. In particular, we show that the frequency of the system drifts away from its theoretical value in the absence of a leader. We implement the algorithm on a cluster of IBM BladeCenter servers running Linux and we experimentally verify that our algorithm outperforms the well-established solution. We also show that the optimal parameter values depend on the network conditions and topology.

Finally, we study synchronization on power networks. By relating the dynamics of power networks to the dynamics of coupled oscillators, we can gain insight into how different network parameters affect performance. We show that the rate of convergence of networks is related to the algebraic connectivity of a state dependent Laplacian which varies with the network power scheduling and line impedances. This provides a novel method to change the voltage stability margins by updating the power scheduling or line impedances. Unfortunately, there exists a decoupling between the market clearing procedure used to dispatch power and the security analysis of the network, that prevents

the direct use of this solution. Furthermore, focusing on voltage stability may generate other types of instabilities such as larger transient oscillations. This motivates the use of a unifying stability measure that can minimize oscillations or maximize voltage stability margins, and can be readily combined with current dispatch mechanisms generating a dynamics-aware optimal power flow formulation.

BIOGRAPHICAL SKETCH

Enrique Mallada received the degree of Telecommunications Engineer from Universidad ORT, Uruguay, in 2005. From 2004 to 2007 he was IT-Specialist at IBM and in 2008 worked as Engineer at the Traffic Engineering Department of ANTEL, the main telecommunications operator in Uruguay. He also was a Teaching and Research Assistant in the Department of Telecommunications at Universidad ORT, and a member of the MATE research group.

He entered the M.S/Ph.D program in the School of Electrical and Computer Engineering at Cornell University in August, 2008. He is a member of the Networks Group, led by Dr. A. Kevin Tang, and member of the FoIE Group. He was recipient of the Organization of American States Scholarship for academic stuides during the 2008-2009 and 2009-2010 academic years and recipient of Jacobs Fellowship of Cornell University in 2011. He coordinated the ISN Seminar throughout the 2010-2011 academic year and was an intern in IBM T. J. Watson Research Center in the summer of 2011. His research interests include dynamical systems, networks, optimization and control.

This thesis is dedicated to my family.

ACKNOWLEDGEMENTS

Looking back at these past years at Cornell, the list of people that in some way helped me get to this point is almost endless. First and foremost, my deepest gratitude is to my advisor A. Kevin Tang. I could not have asked for a better mentor. He has taught me to always look at the big picture, choose my research with a higher goal in mind and leverage previous experience as a competitive edge. I am very grateful for his support and for the freedom he gave me to find my own way, using his meaningful insight as a constant guide. I would also like to thank my committee for their advice and for kindly taking the time to be part of it. Hsiao-Dong Chiang, for teaching me the importance of balancing theory and applications. Steve Strogatz, for sharing his passion for coupled oscillators and teaching me that it is always possible to explain in simple words even the most complex phenomenon. Lang Tong, for teaching me to seek a deep understanding on every problem you face.

During the last five years I had the chance to collaborate with several people which I also owe many gratitude. To Meng Wang and Weiyu Xu from the Networks group, thanks for sharing your research problems, your insights and letting me have fun while working on them. To Li Zhang, Xiaoqiao Meng and Michel Hack from IBM Research, for opening the doors of IBM in many occasions. The work on clock synchronization presented in this thesis would not have been possible without them. Finally, to Randy Freeman from Northwestern University, his thorough examination of my work on heterogeneous frequency oscillators and his brilliant ideas have transformed email exchanges into a very fruitful collaboration.

My deepest thanks also goes to my former advisor in Uruguay, Fernando Paganini. I would not have done a PhD at Cornell if it weren't for him. Thanks

for initiating me into research, for supporting me on every step of this endeavor and being an ever present friend. I have learned from him how to formalize intuition and to always completely understand the simple case before tackling the whole problem. My gratitude also extends to the Grupo Mate's gang: Andrés Ferragut, Martín López, Diego Feijer and Marcos Cardozo. I will never forget the fun, the hard work and willingness to prove that good research can be done in a small country like ours.

I would also like to thank many friends that were in one way or another part of this process. To my life friends from Uruguay, Martín Navia (Palo), Sebastián Cáceres (Nono), Juan Manuel García (Gordo), Paulo Llorach (Adi), Juan Mosca (Largo), Lorena Cal, Helena Muoz, Antonella Vignone, Virginia Reimon and Liliana Barretto, thanks for letting me feel like I never left whenever I see you. I am very lucky of having you as friends. To my latinamerican friends of Ithaca, Andé Velazques, Nicolas Cosentino, Jose Carlos Huguet, Maicol Ochoa, Luis Duque, Margarita Lopez-Uribe, Christine Kraus, Marcelo Aguiar, Juan Alonso, Bernardo Aguilera, Julieta Gallego, Mauricio Bucca, Rosario Donoso, Viviana Sitz, Juliana Rangel, Alejandra Escandón, Felipe Aron and Ronaldo Ilma, for letting me feel a little bit closer to home. Specially, to Silvia Moré and Pancho Arrillaga for being my Ithacan parents. I will never forget all the affection as well as the milanesas! To my officemates Nithin Michael, Chiun Lin Lim, Ilan Shoromony and Alireza Vahid, for countless discussions and many hours of fun.

My family of course, have a special place in this list. Thanks to my parents, for their unconditional love and support, and for giving me the education and moral values that allowed me to get to where I am today. I would also like to thank my sister Maite, for being there to listen when it was much needed and

for being my friend.

Last but not least, my greatest thanks goes to my wife, Josefina. Thanks for your friendship, endless patience and support over the last years. But, more importantly thank for your love. You have given me strength in tough moments and brought happiness to my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

"Synchronicity is an ever present reality for those who have eyes to see."

— Carl Jung

Synchronization is defined in its most general sense as *the coordination of events that allow a system to operate coherently*. It is perhaps one of the most ubiquitous phenomena in nature and science, and its study has widely attracted the attention of researchers in various disciplines such as biology [1–5], chemistry [6, 7] and physics [8, 9]. Perhaps one of the most amazing aspects of synchronization is that it appears to be instrumental in many biological and physical processes. For example, the in-phase synchrony of cells in the sinoatrial node produces the heart contractions responsible for blood circulation [2], the spatial patterns of oscillator chains control the motor patterns of many species [9, 10] and epileptic seizures have been associated with the presence [11,12] or lack [13] of neuron synchronous activity.

In engineering, synchronization has become a fundamental requirement of many distributed applications. Time Division Multiple Access (TDMA) communication systems need to be synchronized in order to coordinate transmissions and decode messages within a network [14,15]. Energy efficient Medium Access Control (MAC) protocols synchronize the sleep periods of the network agents in order to save energy [16–18]. Data fusion of time sensitive measurements in distributed estimation or tracking [19] uses synchronization to minimize estimation error. Also, collaborative transmission systems using space-time coding [20] need synchronization in the transmission instants to properly

function.

However, besides its unusual pervasiveness, the most impressive aspect of synchronization is its ability to emerge in large populations of interconnected (coupled) oscillators without the presence of a specific leader or orchestrator.

1.1 Collective Synchronization

The study of collective synchronization can be traced back to Wiener [21] in 1958. But it was Winfree [22] who formulated the problem as a population of interacting limit-cycle oscillators. In his work, Winfree realized that by assuming weak coupling and making a time scale separation the dimension of the system could be reduced to consider only the phase of each oscillator's orbit. He proposed the following system of N nonlinear differential equations to study synchronization

$$\dot{\phi}_i = \omega_i + \sum_{j \in \mathcal{N}_i} H_{ij}(\phi_i, \phi_j) \quad \forall i \in \{1, ..., N\}. \tag{1.1}$$

Here, ϕ_i is the phase of the ith oscillator, ω_i is the natural frequency of oscillation, H_{ij} denotes the coupling function and $\mathcal{N}_i$ is the set of i's neighbors. Using equation (1.1) in the special case $H_{ij}(\phi_i, \phi_j) = Z(\phi_i)X(\phi_j)$ and $\mathcal{N}_i = \{1, ..., N\}\backslash\{i\}$, plus some additional approximations, Winfree was able to characterize a critical condition for the emergence of collective synchronization.

However, it was not until Kuramoto's work [23] that a theory of collective synchronization started to take shape. Building on Winfree's work, Kuramoto took the phase model provided by assuming weak coupling and used averaging theory to modify equation (1.1) and obtain a coupling that is a function of the

phase difference

$$H_{ij}(\phi_i, \phi_j) = f_{ij}(\phi_j - \phi_i). \tag{1.2}$$

Although equation (1.2) constitutes a significant simplification, the key contribution of Kuramoto was to consider only the first term of the Fourier series of the coupling function, i.e. $f_{ij} = \frac{K}{N} \sin$, which provided analytical tractability. Another closely related line of research comes from assuming pulse-like coupling

$$H_{ij}(\phi_i, \phi_j) = \kappa_{ij}(\phi_i)\delta(\phi_j) \tag{1.3}$$

where δ is a Dirac's delta function. It was first introduced by Peskin [2] in 1975 to study the pacemaker cells of the heart and it has since become a widely used model for many biological processes [24, 25].

Equations (1.1), (1.2) and (1.3) constitute the starting point of different lines of research. By assuming different distributions of ω_i [26–29], taking the continuum limit on the number of oscillators [9, 30, 31] or choosing different communication topologies [32–35] the possible behavior of such a system can be complex and diverse. For example, the intrinsic symmetry of the network can produce multiple limit cycles with relatively fixed phases (phase-locked trajectories) [36], which in many cases can be stable [10]. Also, the heterogeneity in the natural oscillation frequency can lead to incoherence [23] or even chaos [37].

One interesting question, in particular, is whether the coupled oscillators will synchronize (phase lock) in the long run [24, 32, 38–40]. Besides its clear theoretical value, it also has rich applications in practice. Unfortunately, current results present several simplifying assumptions that hinder the potential application of these models in real scenarios. For example, they either restrict to simple topologies, such as complete graph or ring networks, or they assume

zero or bounded delay, homogeneous frequencies, or $\sin$ coupling. This is unsatisfactory as in many applications these assumptions do not hold.

1.2 Synchronization on Information Networks

Keeping consistent time among different nodes in a network is central to many distributed applications on information networks. Their internal clocks are usually not accurate enough and tend to drift apart from each other over time, generating inconsistent time values. This problem is known in engineering and computer science as network clock synchronization. Its solution allows these devices to correct their clocks to match a global reference of time, such as the Universal Coordinated Time (UTC), by performing time measurements through the network. For example, for the Internet, network clock synchronization has been an important subject of research and several different protocols have been proposed [41–47]. These protocols are used for various legacy and emerging applications with diverse precision requirements such as banking transactions, communications, traffic measurement and security protection. In particular, in modern wireless cellular networks, time-sharing protocols need an accuracy of several microseconds to guarantee the efficient use of channel capacity. Another example is the recently announced Google Spanner [48], a globally-distributed database, which depends on globally-synchronized clocks within at most several milliseconds drifts.

The current *de facto* standard for IP networks, NTP [41], is a low-cost, purely software-based solution, yet its accuracy mostly ranges from hundreds of microseconds to several milliseconds, which is often insufficient. On the other

hand, IEEE 1588 (PTP) [43] and IBM CCT [49] give superior performance by achieving sub-microsecond or even nanosecond accuracy (for PTP). However, they are relatively expensive as they require special hardware support to achieve those accuracy levels and may not be fully compatible with legacy cluster systems.

There are three major difficulties that make the problem of network clock synchronization challenging. Firstly, the frequency of hardware clocks is sensitive to temperature, vibrations and interference, and thus constantly varies. Secondly, the latency introduced by OS and network congestion delays results in errors in the time measurements. Thirdly, these time errors can be amplified as they propagate through the network. Thus, most protocols introduce different ways of estimating the frequency mismatch (skew) [50, 51] and measuring the time difference (offset) [52, 53] while maintaining a simple network topology [41, 43].

However, despite the extensive work on this topic [47, 50, 54–57], there are fundamental questions that remain unanswered. In particular, the vast literature on skew estimation [51, 58–60] for clock synchronization suggests that precise estimation of the skew between clocks is needed in order to accurate synchronize them. However, it is not known whether explicit skew estimation is necessary or not.

Furthermore, there is no clear understanding of how network topology and noise affect the synchronization performance. A common practice in the clock synchronization community is to avoid timing loops in the network [41, p. 3] [43, p. 16, s. 6.2]. This is because timing loops are believed to induce instability as stated in [41]: *"Drawing from the experience of the telephone industry,*

which learned such lessons at considerable cost, the subnet topology... must never be allowed to form a loop." Yet to the best of our knowledge there is no theoretical explanation of why and under what conditions loops can produce instability.

1.3 Electric Power Grid: The Largest Synchronized Network Engineered

The american power grid has been regarded as the largest interconnected machine ever engineered by men [61]. Developed for over more than 100 years, it is composed of thousands of interconnected generators that run exactly at the same frequency, and delivers, through its transmission lines, electricity to hundreds of millions of users. In other words, it is the largest synchronous system built by men.

Its stability is one of the major concerns of every utility company. When a blackout occurs, the resulting economic impact can cost between several hundred millions of dollars and a few billion dollars [62–65]. Thus, utility operators are constantly monitoring the network state in order to avoid the various types of instabilities that a power grid might experience. These include, for instance, voltage collapse/instability [66–68], small signal oscillations/instability [69–71] and transient instability [72–74].

Different methods have been developed to assess and prevent each individual stability problem. Voltage stability, for example, can be analyzed using screening and ranking methods [75, 76] and continuation methods that investigate the available transfer capability of the current operating point [77–79].

Small signal oscillations, on the other hand, are locally damped using Power System Stabilizers (PSS) in the exciter control loop [71, 80–86] and globally damped using either power electronics, such as Flexible AC Transmission System (FACTS) devices [70, 87–89], or using Phasor Measurement Unit's (PMU's) information in the PSSs' loop [86]. Finally, transient stability is analyzed using time domain integration [90] or controlling unstable equilibrium point methodology [91, 92].

That said, in order to achieve economical sustainability, utility companies seek to operate the network as efficiently as possible. Thus, every utility company tries to find the best power scheduling that minimizes their specific performance metric (e.g. market welfare, losses, generation cost or voltage magnitudes) subject to physical and operational constraints. This problem is known as the Optimal Power Flow (OPF) and it has a long history in the power systems community, dating back to at least 1962 with the seminal work of Carpentier [93]. Nowadays, the OPF is a fundamental tool for defining prices and arbitrating electricity markets, and many different algorithms have been proposed to solve OPF [94–97].

Unfortunately, there seems to be a gap between performance optimization and stability assessment. For example, in order to perform the stability analysis, it is needed to first fix the power scheduling, which can be either a base case obtained by the OPF or the result of a change in the system (e.g. fault or demand fluctuation), and then studying the stability of the system. While the effect of the scheduling on transient stability is not very clear -as it also depends on the specific fault in consideration, the procedure used to clear it, and the time needed to recover from it (fault clearing time) [98]-, it is certainly critical in volt-

age stability and small signal oscillation studies because the voltage collapse margins and stability of the operating point are directly influenced it.

In fact, many utility companies perform a day ahead detailed stability analysis based on historic records and predictions which is translated into line flow constraints that aim to prevent the OPF from providing a solution that does not meet the predefined stability margins [99–102]. This has two main problems. Firstly, the additional constraints does not have a clear dynamical meaning that can be used to indicate how robust is the current solution. Secondly, it is usually needed to introduce corrections on the scheduling that can generate market inefficiencies.

In summary, this methodology is unable to contemplate the fact that these two problems are intrinsically coupled. This problem has been identified and studied over the last 15 years and several methods have been proposed to include voltage stability constraints in the OPF problem [103–109]. However, adding small signal stability constraints has been a daunting task because it usually requires constraining several (if not all) eigenvalues of the system [102, 110–112]. Furthermore, these procedures can sometimes have undesired outcomes since there is a tradeoff between asymptotic rate of convergence ($\max \Re[\lambda_i]$) and transient amplitude. In other words, improving the asymptotic rate of convergence can increase the amplitude of the oscillations.

1.4 Contributions of This Thesis

Motivated by engineering applications, this thesis focuses on the study of coupled oscillators whose limiting behavior is phase-locked synchronization. That

is, we study a population of oscillators that can lock themselves on a common frequency $\dot{\phi}_i = \omega^*$. We provide a systematic study of synchronization and how it is affected by the different properties of the system, such as coupling, delay, topology and frequency heterogeneity.

The key to the success of our analysis is based on first studying the system with a simplified, yet not trivial, set of assumptions and progressively increasing complexity. By moving from homogeneous frequency towards heterogeneous frequency, we leverage the results of the simpler scenario in order to obtain similar theoretical guarantees in more general instances.

Similarly, we then focus on two specific applications. In both cases, we first find a common ground that allows us to understand these problems using the collective synchronization perspective given by the collective synchronization theory, and then go beyond these idealized models in order to capture the specific challenges and engineering constraints that each application poses.

1.4.1 Coupled Oscillators

In essence, there are three key factors of a system of coupled oscillators that characterize the interaction among oscillators: *coupling*, *delay* and *topology*. For each of them, the existing work has mainly focused on special cases as explained below. In chapter 2, further research is discussed on each of these three factors:

- Topology (*whom to affect*, section 2.2.2): Current results either restrict to complete graph or ring topology for analytical tractability [32], study *local stability* of topology independent solutions over time varying graph [113–

115], or introduce dynamic controllers to achieve synchronization for time-varying uniformly connected graphs [116, 117]. We develop a graph based sufficient condition which can be used to check equilibrium stability for any fixed topology. It also leads to a family of coupling functions that guarantees that the system will reach *global* phase consensus for arbitrary undirected connected graph using only physically meaningful state variables.

- Coupling (*how to affect*, section 2.2.3): The classical Kuramoto model [23] assumes a $\sin()$ coupling function. Our study suggests that certain symmetry and convexity structures should be enough to guarantee global synchronization.
- Delay (*when to affect*, section 2.3): Existing work generally assumes zero delay among oscillators or requires them to be bounded up to a constant fraction of the period [118]. This is clearly unsatisfactory especially if the oscillating frequencies are high. We develop a new framework to study unbounded delays by constructing a non-delayed phase model that is equivalent to the original one. Using this result, we show that wider delay distribution can help reach synchronization.

We then study the effect of heterogeneous natural frequencies in section 2.4. While it is well-known that in-phase synchronization is no longer achievable, we show that by adding an integrator to the dynamics it is possible leverage the results on homogeneous oscillators to re-obtain phase consensus. More precisely, we prove that the same family of coupling functions characterized in the homogenous case achieves global convergence toward the in-phase orbit for almost every initial condition, provided that all these orbits are isolated.

1.4.2 Computer Clock Synchronization

Synchronization of computer clocks is studied in chapter 3. Although tempted to use algorithms like the one proposed in section 2.4, neither of the solutions is satisfactory as they require skew estimation or introduce offset corrections that are undesired. We provide instead a simple algorithm that can compensate the clock skew without any explicit estimation of it. Our algorithm only uses current offset information and an exponential average of the past offsets. Therefore, it neither requires storing long offset history nor does it perform time consuming skew estimation. We analyze the convergence of the algorithm and provide necessary and sufficient conditions for synchronization. The parameter values that guarantee synchronization depend on the interconnection topology, but there is a subset of these that is independent of it and therefore of great practical interest.

We then study the interplay between noise and topology. We show that if the measurements present biased noise, possibly due to queuing delays or forward and backward paths asymmetries, then the system frequency drifts from its theoretical value unless there is a leader[1] in the communication topology. We additionally characterize the effect of topology on the node's mean offset and optimize the system performance by finding a locally optimal set of parameters that minimizes the variance of linear performance metrics. We also discover a rather surprising fact. Even though for some parameter values loops can produce instability, we show that a proper selection of them can guarantee convergence even in the presence of loops. Furthermore, we experimentally demonstrate in section 3.5 that high connectivity between clients, as well as

[1] A node i is a leader of the system if and only if every node j has a path towards i and i has no outgoing link

properly selected parameter values, can actually help reduce the jitter of the synchronization error!

1.4.3 Synchronization on Power Networks

Finally, we concentrate on the study of synchronization on power grids in chapter 4. As discussed in section 1.3, there is an explicit relationship between the network parameters and the system stability which is not easy to characterize. We overcome this difficulty by using our coupled oscillators model from chapter 2 and study the effect of network topology and parameters on the spectral abscissa or asymptotic rate of convergence, i.e. $\max \Re[\lambda_i]$, of the structure preserving power system model introduced in [119]. We first relate $\max \Re[\lambda_i]$ with the algebraic connectivity of a state dependent weighted Laplacian [120] in section 4.2. This evidences the interplay between voltage stability and network topology. Then, in section 4.3, we use the implicit function theorem [121] to explore the dependence of the algebraic connectivity on network parameters. More specifically, we derive how power scheduling and line impedances affect the operating point of the network and predict the net effect of these changes on the algebraic connectivity. With these results, we provide updating rules that can improve the asymptotic rate of convergence $\max \Re[\lambda_i]$ of a power network.

However, these results pose several questions. First, it is not clear whether $\max \Re[\lambda_i]$ is an appropriate metric to measure power grids dynamic performance. In fact, if one focuses entirely on the rate of convergence, the oscillation of the system can increase. We overcome this problem in section 4.4 by using a novel performance metric known as pseudo spectral abscissa, that can bal-

ance transient amplitude and asymptotic convergence rate [122,123]. Using this metric, we propose an optimization framework that imposes voltage and small signal stability constraints on the OPF without explicitly computing and constraining the eigenvalues of the system, and also finds the performance limits of the system.

CHAPTER 2

SYNCHRONIZATION OF COUPLED OSCILLATORS

In this chapter we shall study coupled oscillators, which can be either pulse-coupled or phase-coupled and are derived from assuming *weak* coupling. Although most of the results are presented for phase-coupled oscillators, they can be readily extended for pulse-coupled oscillators (see, e.g., [25,124]). It is worth noting that results in sections 2.2 and 2.4 are independent of the strength of the coupling and therefore do not require the weak coupling assumption

The chapter is organized as follows. We describe pulse-coupled and phase-coupled oscillator models, as well as their common weak coupling approximation, in section 2.1. Using some facts from algebraic graph theory and potential dynamics in section 2.2.1, we present the negative cut instability theorem in section 2.2.2 to check whether an equilibrium is unstable. This leads to Theorem 2.1 in section 2.2.2, which identifies a class of coupling functions that are always synchronized in phase with the system. It is well known that the Kuramoto model produces global synchronization over a complete graph. In section 2.2.3, we demonstrate that a large class of coupling functions, in which the Kuramoto model is a special case, guarantee the instability of most of the limit cycles in a complete graph network. Section 2.3 is devoted to the discussion of the effect of delay. An equivalent non-delayed phase model is constructed whose coupling function is the convolution of the original coupling function and the delay distribution. Using this approach, we show that sometimes more heterogeneous delays among oscillators can help reach synchronization. Finally, we study the effect of heterogenous frequencies in section 2.4. Although in this case in-phase synchronization is no longer for coupled oscillators, we show that by adding

an integrator in the loop together with a linear consensus term, phase consensus is recovered. We also provide a global convergence result under the same conditions of 2.2.2.

2.1 Model Description

We consider two different models of coupled oscillators studied in the literature. The difference between the models arises in the way the oscillators interact, and their dynamics can be quite different. However, when the interactions are weak (weak coupling), both systems behave similarly and share the same approximation. This allows us to study them under a common framework.

Each oscillator is represented by a phase θ_i in the unit circle $\mathbb{S}^1$ which in the absence of coupling moves with constant speed $\dot{\theta}_i = \Omega_i$. Here, $\mathbb{S}^1$ represents the unit circle, or equivalently the interval $[0, 2\pi]$ with 0 and 2π identified ($0 \equiv 2\pi$), and $\Omega_i = \frac{2\pi}{T_i}$ denotes the natural frequency of the oscillation. We will assume that the differences between the natural frequencies are of order ε, i.e. $\Omega_i = \omega + \varepsilon\omega_i$, for some scalar $\varepsilon > 0$, and that the frequency differences ω_i have zero mean ($\sum_{i=1}^{N} \omega_i = 0$).

2.1.1 Pulse-coupled Oscillators

In this model, the interaction between oscillators is performed by pulses. An oscillator j sends out a pulse whenever it crosses zero ($\theta_j = 0$). When oscillator i receives a pulse, it will change its position from θ_i to $\theta_i + \varepsilon\kappa_{ij}(\theta_i)$. The function κ_{ij} represents how the actions of other oscillators affect oscillator i, and the scalar

$\varepsilon > 0$ is a measure of the coupling strength. These jumps can be modeled by a Dirac's delta function δ satisfying $\delta(t) = 0\ \forall t \neq 0$, $\delta(0) = +\infty$, and $\int \delta(s)ds = 1$. The coupled dynamics is represented by

$$\dot{\theta}_i(t) = \Omega_i + \varepsilon \sum_{j \in \mathcal{N}_i} \kappa_{ij}(\theta_i(t))\Omega_j \delta(\theta_j(t - \eta_{ij})), \tag{2.1}$$

where $\eta_{ij} > 0$ is the propagation delay between oscillators i and j ($\eta_{ij} = \eta_{ji}$), and $\mathcal{N}_i$ is the set of i's neighbors. The factor of Ω_j in the sum is needed to keep the size of the jump within $\varepsilon\kappa_{ij}(\theta_i)$. This is because $\theta_j(t)$ behaves like $\Omega_j t$ when crosses zero and therefore the jump produced by $\delta(\theta_j(t))$ is of size $\int \delta(\theta_j(t))dt = \Omega_j^{-1}$ [25].

The coupling function κ_{ij} can be classified based on the qualitative effect it produces in the absence of delay. After one period, if the net effect of the mutual jumps brings a pair of oscillators closer, we call it **attractive** coupling. If the oscillators are brought further apart, it is considered to be **repulsive** coupling. The former can be achieved, for instance, if $\kappa_{ij}(\theta) \leq 0$ for $\theta \in [0, \pi)$ and $\kappa_{ij}(\theta) \geq 0$ for $\theta \in [\pi, 2\pi)$. See Figure 2.1 for an illustration of an attractive coupling κ_{ij} and its effect on the relative phases.

This pulse-like interaction between oscillators was first introduced by Peskin [2] in 1975 as a model of the pacemaker cells of the heart, although its canonic form did not appear in the literature until 1999 [25]. In general, when the number of oscillators is large, there are several different limit cycles besides the in-phase synchronization and many of them can be stable [10].

The question of whether this system can collectively achieve in-phase synchronization was answered for the complete graph case and zero delay by Mirollo and Strogatz in 1990 [24]. They showed that if $\kappa_{ij}(\theta)$ is strictly increasing

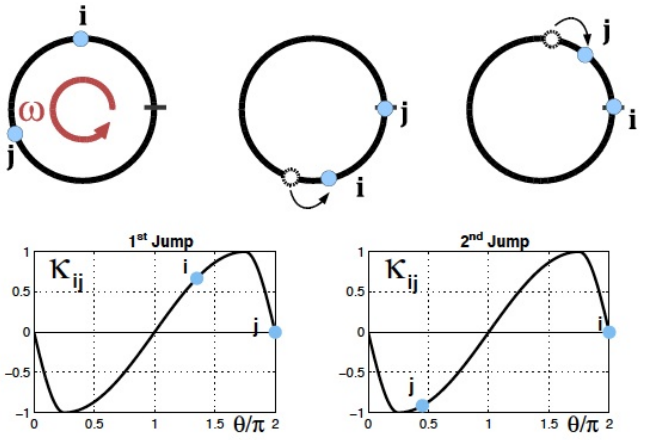


Figure 2.1: Pulse-coupled oscillators with attractive coupling.

on $(0, 2\pi)$ with a discontinuity in 0 (which resembles attractive coupling), then for almost every initial condition, the system can synchronize in phase in the long run.

The two main assumptions of [24] are all to all communication and zero delay. Whether in-phase synchronization can be achieved for arbitrary graphs has been an open problem for over twenty years. On the other hand, when delay among oscillators is introduced the analysis becomes intractable. Even for the case of two oscillators, the number of possibilities to be considered is large [125, 126].

2.1.2 Phase-coupled Oscillators

In the model of phase-coupled oscillators, the interaction between neighboring oscillators i and $j \in \mathcal{N}_i$ is modeled by change of the oscillating speeds. Although in general the speed change can be a function of both phases (θ_i, θ_j), we concentrate on the case where the speed change is a function of the phase differences

$f_{ij}(\phi_j(t-\eta_{ij})-\phi_i(t))$. Thus, since the net speed change of oscillator i amounts to the sum of the effects of its neighbors, the full dynamics is described by

$$\dot{\phi}_i(t) = \Omega_i + \varepsilon \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j(t-\eta_{ij})-\phi_i(t)). \tag{2.2}$$

The function f_{ij} is usually called coupling function, and as before η_{ij} represents delay and $\mathcal{N}_i$ is the set of neighbors of i.

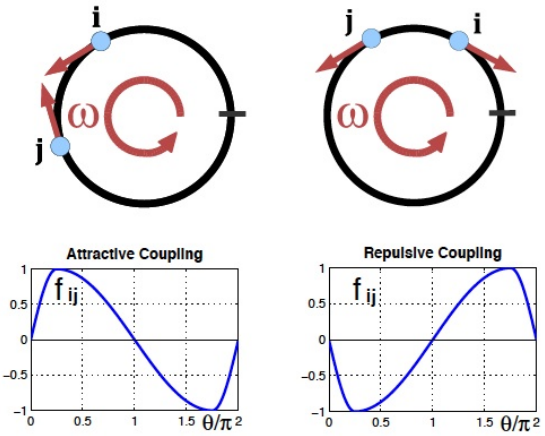


Figure 2.2: Phase-coupled oscillators with attractive and repulsive coupling.

A similar definition for attractive and repulsive couplings can be done in this model. We say that the coupling function f_{ij} is **attractive** if, without delays, the change in speeds brings oscillators closer, and **repulsive** if they are brought apart. Figure 2.2 shows typical attractive and repulsive coupling functions where arrows represent the speed change produced by the other oscillator; if the pointing direction is counter clockwise, the oscillator speeds up, and otherwise it slows down.

When $f_{ij} = \frac{K}{N}\sin()$, $K > 0$ (attractive coupling), this model is known as the classical Kuramoto model [127]. Intensive research has been conducted on this model, but convergence results are usually limited to cases with all to all

coupling ($\mathcal{N}_i = \mathcal{N}\backslash\{i\}$, i.e., complete graph topology) and no delay ($\eta_{ij} = 0$), see e.g. [32, 128], or to some regions of the state space [118].

2.1.3 Weak Coupling Approximation

We now concentrate on the regime in which the coupling strength of both models is weak, i.e. $1 \gg \varepsilon > 0$. For pulse-coupled oscillators, this implies that the effect of the jumps originated by each neighbor can be approximated by their average [124]. For phase-coupled oscillators, it implies that to the first order $\phi_i(t - \eta_{ij})$ is well approximated by $\phi_i(t) - \omega\eta_{ij}$.

The effect of these approximations allows us to completely capture the behavior of both systems using the following equation

$$\dot{\phi}_i = \varepsilon\omega_i + \varepsilon \sum_{j\in\mathcal{N}_i} f_{ij}(\phi_j - \phi_i - \psi_{ij}). \tag{2.3}$$

where we know that ϕ_i is the phase of a rotating frame of speed ω and we only keep track of the slow time scale of order $\frac{1}{\varepsilon}$. Furthermore, since ε multiplies both terms on the right hand side of (2.3), we will drop it without loss of generality. That is, we will consider

$$\dot{\phi}_i = \omega_i + \sum_{j\in\mathcal{N}_i} f_{ij}(\phi_j - \phi_i - \psi_{ij}). \tag{2.4}$$

For pulse-coupled oscillators, the coupling function is given by

$$f_{ij}(\theta) = \frac{\omega}{2\pi}\kappa_{ij}(-\theta), \tag{2.5}$$

and the phase lag $\psi_{ij} = \omega\eta_{ij}$ represents the distance that the phase of oscillator i can travel along the unit circle during the delay time η_{ij}. Equation (2.5) also

shows that the attractive/repulsive coupling classification of both models is in fact equivalent, since in order to produce the same effect κ_{ij} and f_{ij} should be mirrored, as illustrated in Figure 2.1 and Figure 2.2.

Equation (2.4) captures the relative change of the phases and therefore any solution to (2.4) can be immediately translated to either (2.1) or (2.2) by rescaling time and adding ωt. For example, if ϕ^* is an equilibrium of (2.4), by adding ωt, we obtain a limit cycle in the previous models. Besides the delay interpretation for ψ_{ij}, (2.4) is also known as a system of coupled oscillators with *frustration*, see e.g. [129].

From now on we will concentrate on (2.4) with the understanding that any convergence result derived will be immediately true for the original models in the weak coupling limit. We are interested in the attracting properties of phase-locked invariant orbits within $\mathcal{T}^N$, which can be represented by

$$\phi(t) = \omega^* t \mathbf{1}_N + \phi^*, \tag{2.6}$$

where $\mathbf{1}_N = (1, \ldots, 1)^T \in \mathcal{T}^N$, and ϕ^* and ω^* are solutions to

$$\omega^* = \omega_i + \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j^* - \phi_i^* - \psi_{ij}), \forall i. \tag{2.7}$$

Whenever the system reaches one of these orbits, we say that it is synchronized or phase-locked. If all the elements of ϕ^* are equal, we say the system is synchronized **in-phase** or that it is **in-phase** locked.

Moreover, if ϕ^* is an equilibrium of (2.4), any solution of the form $\phi^* + \lambda \mathbf{1}_N$, with $\lambda \in \mathbb{R}$, is also an equilibrium that identifies the same limit cycle on the original system. Therefore, two equilibria $\phi^{1,*}$ and $\phi^{2,*}$ will be considered to be equivalent, if both identify the same orbit, or equivalently, if both belong to the

same connected set of equilibria

$$E_{\phi^*} := \{\phi \in \mathcal{T}^N | \phi = \phi^* + \lambda \mathbf{1}_N, \lambda \in \mathbb{R}\}. \tag{2.8}$$

In the next two sections (section 2.2 and section 2.3) we will assume that the natural frequency is homogeneous among the population of oscillators, i.e. $\Omega_i = \omega \ \forall i$ and $\omega_i = 0$.

2.2 Effect of Topology and Coupling

In this section, we concentrate on the class of coupling functions f_{ij} that are symmetric ($f_{ij} = f_{ji} \ \forall ij$), odd ($f_{ij}(-\theta) = -f_{ij}(\theta)$) and continuously differentiable. We also assume that there is no delay within the network ($\psi_{ij} = 0 \ \forall ij$). Thus, (2.4) reduces to

$$\dot{\phi}_i = \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j - \phi_i). \tag{2.9}$$

In the rest of this section, we progressively show how with some extra conditions on f_{ij} we can guarantee in-phase synchronization for arbitrary undirected graphs. Since the network can have many other phase-locked trajectories besides the in-phase one, our target is an **almost global stability** result [130], meaning that the set of initial conditions that does not eventually lock in-phase has zero measure. Later we show how most of the phase-locked solutions that appear on a complete graph are unstable under some general conditions on the structure of the coupling function.

2.2.1 Preliminaries

We now introduce some prerequisites used in our later analysis.

Algebraic Graph Theory

We start by reviewing basic definitions and properties from graph theory [131, 132] that are used in this chapter. Let G be the connectivity graph that describes the coupling configuration. We use $V(G)$ and $E(G)$ to denote the set of vertices (i or j) and undirected edges (e) of G. An undirected graph G can be directed by giving a specific orientation σ to the elements in the set $E(G)$. That is, for any given edge $e \in E(G)$, we designate one of the vertices to be the *head* and the other to be the *tail* giving G^σ.

Although in the definitions that follow we need to give the graph G a given orientation σ, the underlying connectivity graph of the system is assumed to be **undirected**. This is not a problem as the properties used here are independent of a particular orientation σ and therefore they are properties of the undirected graph G. Thus, to simplify notation we drop the superscript σ from G^σ with the understanding that G is now an induced directed graph with some fixed, but arbitrarily chosen, orientation.

We use $P = (V^-, V^+)$ to denote a partition of the vertex set $V(G)$ such that $V(G) = V^- \cup V^+$ and $V^- \cap V^+ = \emptyset$. The cut $C(P)$ associated with P, or equivalently $C(V^-, V^+)$, is defined as $C(P) := \{ij \in E(G) | i \in V^-, j \in V^+, \text{ or vice versa.}\}$. Each partition can be associated with a vector column c_P where $c_P(e) = 1$ if e goes form V^- to V^+, $c_P(e) = -1$ if e goes form V^+ to V^- and $c_P(e) = 0$ if e stays within either set.

There are several matrices associated with the oriented graph G that embed information about its topology. However, the one with most significance to this work is the *oriented incidence matrix* $B \in \mathbb{R}^{|V(G)|\times|E(G)|}$ where $B(i,e) = 1$ if i is the head of e, $B(i,e) = -1$ if i is the tail of e and $B(i,e) = 0$ otherwise.

Potential Dynamics

We now describe how our assumptions on f_{ij} not only simplify the dynamics considerably but also allow us to use the graph theory properties introduced in Section 2.2.1 for a deeper understanding of (2.4).

While f_{ij} being continuously differentiable is a standard assumption to study local stability and it is sufficient to apply LaSalle's invariance principle [133], the symmetry and odd assumptions have a stronger effect on the dynamics.

For example, under these assumptions the system (2.9) can be compactly rewritten in a vector form as

$$\dot{\phi} = -BF(B^T\phi) \tag{2.10}$$

where B is the adjacency matrix defined in Section 2.2.1 and the map $F : \mathcal{E}(G) \to \mathcal{E}(G)$ is

$$F(y) = (f_{ij}(y_{ij}))_{ij\in E(G)}.$$

This new representation has several properties. First, from the properties of B one can easily show that (2.7) can only hold with $\omega^* = 0$ for arbitrary graphs [38] (since $N\omega^* = \omega^*\mathbf{1}_N^T\mathbf{1}_N = -\mathbf{1}_N^T BF(B^T\phi) = 0$), which implies that every

phase-locked solution is an equilibrium of (2.9) and that every limit cycle of the original system (2.4) can be represented by some E^*_ϕ on (2.9).

However, the most interesting consequence of (2.10) comes from interpreting $F(y)$ as the gradient of a potential function

$$V(y) = \sum_{ij \in E(G)} \int_0^{y_{ij}} f_{ij}(s) ds.$$

Then, by evaluating it at $B^T\phi$, (2.10) becomes a gradient descent law for $V(B^T\phi)$, i.e.,

$$\dot{\phi} = -BF(B^T\phi) = -B\nabla V(B^T\phi) = -\nabla(V \circ B^T)(\phi).$$

This makes $V(B^T\phi)$ a natural Lyapunov function candidate since

$$\dot{V}(B^T\phi) = \langle \nabla(V \circ B^T)(\phi), \dot{\phi} \rangle = -\left|\nabla(V \circ B^T)(\phi)\right|^2 = -\left|\dot{\phi}\right|^2 \leq 0. \tag{2.11}$$

Furthermore, since the trajectories of (2.10) are constrained into the N-dimensional torus $\mathcal{T}^N$, which is compact, $V(B^T\phi)$ satisfies the hypothesis of LaSalle's invariance principle (Theorem 4.4 [133]), i.e. there is a compact positively invariant set, $\mathcal{T}^N$ and a function $V \circ B^T : \mathcal{T}^N \to \mathbb{R}$ that decreases along the trajectories $\phi(t)$. Therefore, for every initial condition, the trajectory converges to the largest invariant set M within $\{\dot{V}(B^T\phi) \equiv 0\}$ which is the equilibria set $E = \{\phi \in \mathcal{T}^N | \dot{\phi} \equiv 0\} = \bigcup_{\phi^*} E_{\phi^*}$.

Remark 2.1. *The fact that symmetric and odd coupling induces potential dynamics is well know in the physics community [134]. However, it has also been rediscovered in the control community [39] for the specific case of sine coupling. Clearly, this is not enough to show almost global stability, since it is possible to have other stable phase-locked equilibrium sets besides the in-phase set. However, if we are able to show that all the non-in-phase equilibria are unstable, then almost global stability follows. That is the focus of the next section.*

2.2.2 Negative Cut Instability Condition

We now present the main results of this section. Our technique can be viewed as a generalization of [32]. By means of algebraic graph theory, we provide a better stability analysis of the equilibria under a more general framework. We also use the new stability results to characterize f_{ij} that guarantees almost global stability.

Local Stability Analysis

In this section, we develop the graph theory based tools to characterize the stability of each equilibrium. We will show that given an equilibrium ϕ^* of the system (2.10), with connectivity graph G and f_{ij} as described in this section. If there is a cut $C(P)$ such that the sum

$$\sum_{ij \in C(P)} f'_{ij}(\phi^*_j - \phi^*_i) < 0, \tag{2.12}$$

the equilibrium ϕ^* is **unstable**.

Consider first an equilibrium point ϕ^*. Then, the first order approximation of (2.10) around ϕ^* is

$$\delta\dot{\phi} = -L(w(\phi^*))\delta\phi$$

were $\delta\phi = \phi - \phi^*$ is the incremental phase variable, and

$$L(w(\phi^*)) := B\mathrm{diag}[w(\phi^*)]B^T \tag{2.13}$$

is a state dependent Laplacian matrix $L(w(\phi^*)) \in \mathbb{R}^{|V(G)|\times|V(G)|}$, $(w(\phi^*))_{ij} = f'_{ij}(\phi^*_j - \phi^*_i)$ and $\mathrm{diag}[w(\phi^*)] := \frac{\partial}{\partial y}F(B^T\phi^*) \in \mathbb{R}^{|E(G)|\times|E(G)|}$ is the Jacobian of $F(y)$ evaluated at $B^T\phi^*$.

Now let $A = -L(w(\phi^*))$ and consider the linear system $\delta\dot{\phi} = A\delta\phi$. Although it is possible to numerically calculate the eigenvalues of A given ϕ^* to study the stability, here we use the special structure of A to provide a sufficient condition for instability that has nice graph theoretical interpretations.

Since A is symmetric, it is straight forward to check that A has at least one positive eigenvalue, i.e. ϕ^* is unstable, if and only if $x^T Ax > 0$. Now, given any partition $P = (V^-, V^+)$, consider the associated vector c_P, define x_P such that $x_i = \frac{1}{2}$ if $i \in V^+$ and $x_i = -\frac{1}{2}$ if $i \in V^-$. Then it follows from the definition of B that $c_P = B^T x_P$ which implies that

$$-x_P^T A x_P = c_P^T \mathrm{diag}[w(\phi^*)] c_P = \sum_{ij \in C(P)} f'_{ij}(\phi_j^* - \phi_i^*).$$

Therefore, when condition (2.12) holds, $A = -L(w(phi^*))$ has at least one eigenvalue whose real part is **positive**.

Remark 2.2. *Equation* (2.12) *provides a* **sufficient** *condition for instability; it is not clear what happens when* (2.12) *does not hold. However, it gives a graph-theoretical interpretation that can be used to provide stability results for general topologies. That is, if the* **minimum cut cost** *is negative, the equilibrium is unstable.*

Remark 2.3. *Since the weights of the graph $f'_{ij}(\phi_j^* - \phi_i^*)$ are functions of the phase difference,* (2.12) *holds for any equilibria of the form $\phi^* + \lambda \mathbf{1}_N$. Thus, the result holds for the whole set E_{ϕ^*} defined in* (2.8).

When (2.12) is specialized to $P = (\{i\}, V(G)\backslash\{i\})$ and $f_{ij}(\theta) = \sin(\theta)$, it reduces to the instability condition in Lemma 2.3 of [32]; i.e.,

$$\sum_{j \in \mathcal{N}_i} cos(\phi_j^* - \phi_i^*) < 0. \tag{2.14}$$

However, (2.12) has a broader applicability spectrum as the following example shows.

Example 2.1. *Consider a six oscillators network as in Figure 2.3, where each node is linked to its four closest neighbors and* $f_{ij}(\theta) = \sin(\theta)$. *Then, by symmetry, it is easy to verify that*

$$\phi^* = \left[0, \frac{\pi}{3}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{5\pi}{3}\right]^T \tag{2.15}$$

is an equilibrium of (2.9).

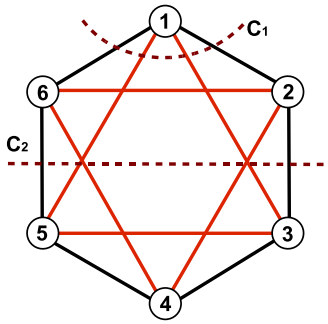


Figure 2.3: The network of six oscillators (Example 4)

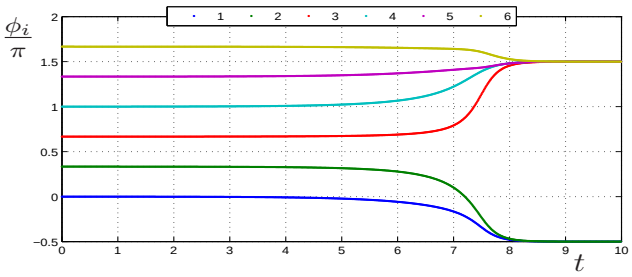


Figure 2.4: Unstable equilibrium ϕ^*. Initial condition $\phi_0 = \phi^* + \delta\phi$

We first study the stability of ϕ^* *using* (2.14) *as in [32]. By substituting* (2.15) *in* $\cos(\phi_j^* - \phi_i^*)$ $\forall ij \in E(G)$ *we find that the edge weights can only take two values:*

$$\cos(\phi_j^* - \phi_i^*) = \begin{cases} \cos(\frac{\pi}{3}) = \frac{1}{2}, & \text{if } j = i \pm 1 \mod 6 \\ \cos(\frac{2\pi}{3}) = -\frac{1}{2}, & \text{if } j = i \pm 2 \mod 6 \end{cases}$$

Then, since any cut that isolates one node from the rest (like $C_1 = C(\{1\}, V(G)\backslash\{1\})$ in Figure 2.3) will always have two edges of each type, their sum is **zero**. *Therefore,* (2.14) *cannot be used to determine stability.*

If we now use condition (2.12) *instead, we are allowed to explore a wider variety of cuts that can potentially have smaller costs. In fact, if instead of C_1 we sum over $C_2 = C(\{1,2,6\},\{3,4,5\})$, we obtain,*

$$\sum_{ij\in C_2} \cos(\phi_j^* - \phi_i^*) = -1 < 0,$$

which implies that ϕ^* is unstable.

Figure 2.4 verifies the equilibrium instability. By starting with an initial condition $\phi_0 = \phi^* + \delta\phi$ close to the equilibrium ϕ^*, we can see how the system slowly starts to move away from ϕ^* towards a **stable** *equilibrium set.*

Furthermore, we can study the whole family of non-isolated equilibria given by

$$\phi^* = \left[\varepsilon_1, \frac{\pi}{3} + \varepsilon_2, \frac{2\pi}{3} + \varepsilon_3, \pi + \varepsilon_1, \frac{4\pi}{3} + \varepsilon_2, \frac{5\pi}{3} + \varepsilon_3\right]^T \tag{2.16}$$

where $\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \mathbb{R}$, which due to Remark 2.3, we can reduce (2.16) *to*

$$\phi^* = \left[0, \frac{\pi}{3} + \lambda_1, \frac{2\pi}{3} + \lambda_2, \pi, \frac{4\pi}{3} + \lambda_1, \frac{5\pi}{3} + \lambda_2\right]^T \tag{2.17}$$

with $\lambda_1 = \varepsilon_2 - \varepsilon_1$ and $\lambda_2 = \varepsilon_3 - \varepsilon_1$.

Instead of focusing on only one cut, here we compute the minimum cut value (2.12) *over the 31 possible cuts, i.e. $C^*(\lambda_1, \lambda_2) := \min_P \sum_{ij\in C(P)} f'_{ij}(\phi_j(\lambda_1,\lambda_2)^* - \phi_i^*(\lambda_1,\lambda_2))$. Figure 2.5 shows the value of $C^*(\lambda_1, \lambda_2)$ for $\lambda_i \in [-\pi, \pi]$. Since $C^*(\lambda_1, \lambda_2)$ is 2π-periodic on each variable and its value is negative for every $\lambda_1, \lambda_2 \in [-\pi, \pi]$, the family of equilibria* (2.17) *(and consequently* (2.16)*) is unstable.*

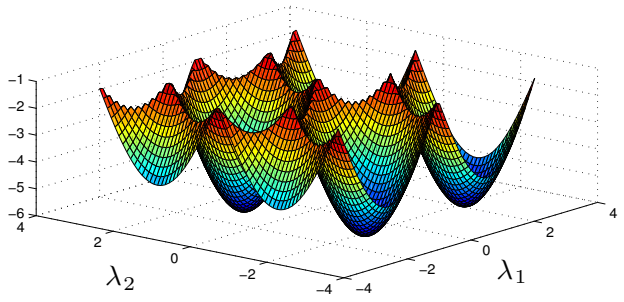


Figure 2.5: Minimum cut value $C^*(\lambda_1, \lambda_2)$ showing that the equilibria (2.16) are unstable

Almost Global Stability

Condition (2.12) also provides insight on which class of coupling functions can potentially give us almost global convergence to the in-phase equilibrium set $E_{\mathbf{1}_N}$. If it is possible to find some f_{ij} with $f'_{ij}(0) > 0$, such that for any non-in-phase equilibrium ϕ^*, there is a cut C with $\sum_{ij \in C} f'_{ij}(\phi^*_j - \phi^*_i) < 0$, then the in-phase equilibrium set will be almost globally stable [10]. The main difficulty is that for general f_{ij} and arbitrary network G, it is not easy to locate every phase-locked equilibria and it is therefore hard to know in what region of the domain of f_{ij} the slope should be negative.

We now concentrate on the one-parameter family of functions $\mathcal{F}_b$.

Definition 2.1. *$f_{ij}(\theta; b)$ is a member of $\mathcal{F}_b$ in and only if:*

- *Symmetric ($f_{ij} = f_{ji}\ \forall ij$), odd ($f_{ij}(-\theta) = -f_{ij}(\theta)$) and continuously differentiable ($f_{ij} \in C^1$)*
- $f'_{ij}(\theta; b) > 0, \forall \theta \in (0, b) \cup (2\pi - b, 2\pi)$, *and*
- $f'_{ij}(\theta; b) < 0, \forall \theta \in (b, 2\pi - b)$.

See Figure 2.6 for an illustration with $b = \frac{\pi}{2}$ and $\frac{\pi}{6}$. Also note that this definition implies that if $f_{ij}(\theta; b) \in \mathcal{F}_b$, the coupling is attractive and $f_{ij}(\theta; b) > 0$ $\forall \theta \in (0, \pi)$. This last property will be used later. We also assume the graph G to be **connected**.

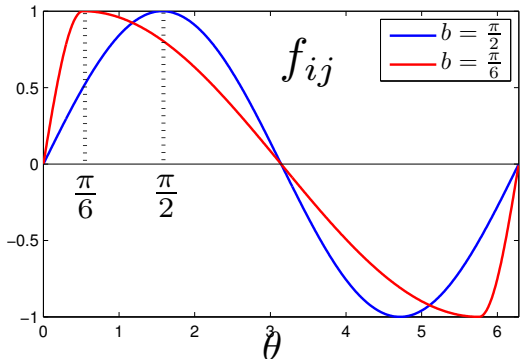


Figure 2.6: Coupling function $f_{ij} \in \mathcal{F}_b$ for $b = \frac{\pi}{2}$ and $b = \frac{\pi}{6}$

In order to obtain almost global stability we need b to be small. However, since the equilibria position is not known a priori, it is not clear how small b should be or if there is any $b > 0$ such that all nontrivial equilibria are unstable. We therefore need to first estimate the region of the state space that contains every non-trivial phase-locked solution.

Let I be a compact connected subset of $\mathbb{S}^1$ and let $l(I)$ be its length, e.g., if $I = \mathbb{S}^1$ then $l(I) = 2\pi$. For any $S \subset V(G)$ and $\phi \in \mathcal{T}^N$, define $d(\phi, S)$ as the length of the smallest interval I such that $\phi_i \in I$ $\forall i \in S$, i.e.

$$d(\phi, S) = l(I^*) = \min_{I : \phi_i \in I,\, \forall i \in S} l(I).$$

Using this metric, together with the aid of Theorem 2.6 of [38] we can identify two very insightful properties of the family $\mathcal{F}_b$ whenever the graph G is connected.

Lemma 2.1. *If ϕ^* is an equilibrium point of* (2.10) *with $d(\phi^*, V(G)) \leq \pi$, then either ϕ^* is an in-phase equilibrium, i.e. $\phi^* = \lambda \mathbf{1}_N$ for $\lambda \in \mathbb{R}$, or has a cut C with $f'_{ij}(\phi^*_j - \phi^*_i) < 0$ $\forall ij \in C$.*

Proof. Since $d(\phi^*, V(G)) \leq \pi$, all the phases are contained in a half circle and for the oscillator with smallest phase i_0, all the phase differences $(\phi^*_j - \phi^*_{i_0}) \in [0, \pi]$. However, since $f_{ij}(\cdot; b) \in \mathcal{F}_b$ implies $f_{ij}(\theta; b) \geq 0$ $\forall \theta \in [0, \pi]$ with equality only for $\theta \in \{0, \pi\}$, $\dot{\phi}^*_{i_0} = \sum_{j \in \mathcal{N}_{i_0}} f_{ij}(\phi^*_j - \phi^*_{i_0}) = 0$ can only hold if $\phi^*_j - \phi^*_{i_0} \in \{0, \pi\}$ $\forall j \in \mathcal{N}_{i_0}$. Now let $V^- = \{i \in V(G) : d(\phi^*, \{i, i_0\}) = 0\}$ and $V^+ = V(G) \backslash V^-$. If $V^- = V(G)$, then ϕ^* is an in-phase equilibrium. Otherwise, $\forall ij \in C(V^-, V^+)$, $f'_{ij}(\phi^*_j - \phi^*_i) = f'_{ij}(\pi) < 0$. □

We are now ready to establish a bound on the value of b that guarantees the instability of the non-in-phase equilibria.

Lemma 2.2. *Consider $f_{ij}(\cdot; b) \in \mathcal{F}_b$ $\forall ij \in E(G)$ and arbitrary connected (undirected) graph G. Then for any $b \leq \frac{\pi}{N-1}$ and non-in-phase equilibrium ϕ^*, there is a cut C with $f'_{ij}(\phi^*_j - \phi^*_i; b) < 0, \forall ij \in C$*

Proof. Suppose there is a non-in-phase equilibrium ϕ^* for which no such cut C exists. Let $V_0^- = \{i_0\}$ and $V_0^+ = V(G) \backslash \{i_0\}$ be a partition of $V(G)$ for some arbitrary node i_0.

Since such C does not exist, there is some edge $i_0 j_1 \in C(V_0^-, V_0^+)$, with $j_1 \in V_0^+$, such that $f'_{i_0 j_1}(\phi^*_{j_1} - \phi^*_{i_0}; b) \geq 0$. Move j_1 from one side to the other of the partition by defining $V_1^- := V_0^- \cup \{j_1\}$ and $V_1^+ := V_0^+ \backslash \{j_1\}$. Now since $f'_{i_0 j_1}(\phi^*_{j_1} - \phi^*_{i_0}; b) \geq 0$, then

$$d(\phi^*, V_1^-) \leq b.$$

In other words, both phases should be within a distance smaller than b.

Now repeat the argument k times. At the k^{th} iteration, given V_{k-1}^-, V_{k-1}^+, again we can find some $i_{k-1} \in V_{k-1}^-$, $j_k \in V_{k-1}^+$ such that $i_{k-1}j_k \in C(V_{k-1}^-, V_{k-1}^+)$ and $f'_{i_{k-1}j_k}(\phi^*_{j_k} - \phi^*_{i_{k-1}}; b) \geq 0$. Also, since at each step $d(\phi^*, \{i_{k-1}, j_k\}) \leq b$,

$$d(\phi^*, V_k^-) \leq b + d(\phi^*, V_{k-1}^-).$$

Thus by solving the recursion we get: $d(\phi^*, V_k^-) \leq kb$.

After $N-1$ iterations we have $V_{N-1}^- = V(G)$ and $d(\phi^*, V(G)) \leq (N-1)b$. Therefore, since $b \leq \frac{\pi}{N-1}$, we obtain

$$d(\phi^*, V(G)) \leq (N-1)\frac{\pi}{N-1} = \pi.$$

Then, by Lemma 2.1 ϕ^* is either an in-phase equilibrium or there is a cut C with $f'_{ij}(\phi^*_j - \phi^*_i) < 0 \; \forall ij \in C$. Either case gives a contradiction to assuming that ϕ^* is a non-in-phase equilibrium and C does not exist. Therefore, for any non-in-phase ϕ^* and $b \leq \frac{\pi}{N-1}$, we can always find a cut C with $f'_{ij}(\phi^*_j - \phi^*_i; b) < 0, \forall ij \in C$. □

Lemma 2.2 allows us to use our cut condition (2.12) on every non-in-phase equilibrium. Thus, since (2.10) is a potential dynamics (c.f. section 2.2.1), from every initial condition the system converges to the set of equilibria E. But when $b \leq \frac{\pi}{N-1}$ the only stable equilibrium set inside E is the in-phase set $E_{\mathbf{1}_N}$. Thus, $E_{\mathbf{1}_N}$ set is globally asymptotically stable. We have summarized this result in the following Theorem.

Theorem 2.1 (Almost global stability)**.** *Consider $f_{ij}(\theta; b) \in \mathcal{F}_b$ and an arbitrary connected graph G. Then, if $b \leq \frac{\pi}{N-1}$, the in-phase equilibrium set $E_{\mathbf{1}_N}$ is* **almost globally asymptotically stable**.

This result provides a sufficient condition for almost global asymptotic stability to the in-phase equilibrium set $E_{\mathbf{1}_N}$. Although found independently, the same condition was proposed for a specific piecewise linear f_{ij} in [135]. Here we extend [135] in many aspects. For example, instead of assuming equal coupling for every edge, our condition describes a large family of coupling functions $\mathcal{F}_b$ where each f_{ij} can be taken independently from $\mathcal{F}_b$. Also, in [135] the construction of $f_{ij}(\theta)$ assumes a discontinuity on the derivative at $\theta = b$. This can pose a problem if the equilibrium ϕ^* happens to have phase differences $\phi_j^* - \phi_i^* = b$. Here we do not have such problem as f_{ij} is continuously differentiable.

The condition $b \leq \frac{\pi}{N-1}$ implies that, when N is large, f_{ij} should be decreasing in most of it domain. Using (2.5) this implies that κ_{ij} should be increasing within the region $(b, 2\pi - b)$, which is similar to the condition on [24] and equivalent when $b \to 0$. Thus, Theorem 2.1 confirms the conjecture of [24] by extending their result to arbitrary topologies and a more realistic continuous κ_{ij} for the system (2.1) in the weak coupling limit.

2.2.3 Complete Graph Topology with a Class of Coupling Functions

In this subsection, we investigate how conservative the value of b found in section 2.2.2 is for the complete graph topology. We are motivated by the results of [32] where it is shown that $f(\theta) = \sin(\theta)$ ($b = \frac{\pi}{2}$) with complete graph topology ensures almost global synchronization.

Since for general f it is not easy to characterize all the possible equilibria of

the system, we study the stability of the equilibria that appear due to the equivalence of (2.10) with respect to the action group $S_N \times T^1$, where S_N is the group of permutations of the N coordinates and $T^1 = [0, 2\pi)$ represents the group action of phase shift of all the coordinates, i.e. the action of $\delta \in T^1$ is $\phi_i \mapsto \phi_i + \delta$ $\forall i$. We refer the readers to [36] and [38] for a detailed study of the effect of this property.

These equilibria are characterized by the isotropy subgroups Γ of $S_N \times T^1$ that keep them fixed, i.e., $\gamma\phi^* = \phi^*$ $\forall \gamma \in \Gamma$. In [36] it was shown that this isotropy subgroup takes the form of

$$(S_{k_0} \times S_{k_1} \times \cdots \times S_{k_{l_B-1}})^m \rtimes Z_m$$

where k_i and m are positive integers such that $(k_0+k_1+\cdots+k_{l_B-1})m = N$, S_j is the permutation subgroup of S_N of j-many coordinates and Z_m is the cyclic group with action $\phi_i \mapsto \phi_i + \frac{2\pi}{m}$. The semiproduct $\rtimes$ represents the fact that Z_m does not commute with the other subgroups. In other words, each equilibria with isotropy $(S_{k_0} \times S_{k_1} \times \cdots \times S_{k_{l_B-1}})^m \rtimes Z_m$ is conformed by l_B shifted constellations C_l ($l \in \{0, 1, \ldots l_B - 1\}$) of m evenly distributed blocks, with k_l oscillators per block. We use δ_l to denote the phase shift between constellation C_0 and C_l. See Figure 2.7 for examples of these types of equilibria.

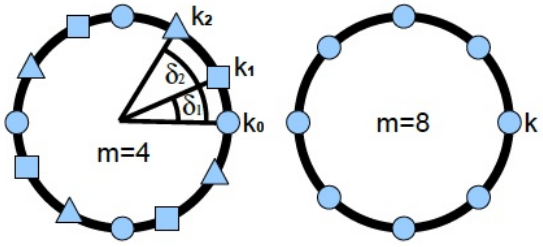


Figure 2.7: Equilibria with isotropy $(S_{k_0} \times S_{k_1} \times S_{k_2})^4 \rtimes Z_4$ (left) and $(S_k)^8 \rtimes Z_8$ (right)

Here we will show that under mild assumptions on f and for $b = \frac{\pi}{2}$, most of the equilibria found with these characteristics are unstable. We first study all the equilibria with m even. In this case there is a special property that can be exploited.

That is, when $f \in \mathcal{F}_{\frac{\pi}{2}}$ such that f is even around $\frac{\pi}{2}$, we have

$$\begin{aligned} g_m(\delta) &:= \sum_{j=0}^{m-1} f(\frac{2\pi}{m}j + \delta) \\ &= \sum_{j=0}^{m/2-1} f(\frac{2\pi}{m}j + \delta) + f(\pi + \frac{2\pi}{m}j + \delta) \\ &= \sum_{j=0}^{m/2-1} f(\frac{2\pi}{m}j + \delta) + f((\frac{3\pi}{2} + \frac{2\pi}{m}j + \delta) - \frac{\pi}{2}) \\ &= \sum_{j=0}^{m/2-1} f(\frac{2\pi}{m}j + \delta) + f(-(\frac{2\pi}{m}j + \delta)) \\ &= \sum_{j=0}^{m/2-1} f(\frac{2\pi}{m}j + \delta) - f(\frac{2\pi}{m}j + \delta) = 0 \end{aligned} \tag{2.18}$$

where the third step comes from f being even around $\pi/2$ and 2π-periodic, and the fourth from f being odd.

Having $g_m(\delta) = 0$ is the key to prove the instability of every equilibria with even m. It essentially states that the aggregate effect of one constellation C_l on any oscillator $j \in V(G)\backslash C_l$ is zero when m is even, and therefore any perturbation that maintains C_l has null effect on j. This is shown in the next proposition.

Theorem 2.2 (Instability for even m)**.** *Given an equilibrium ϕ^* with isotropy $(S_{k_1} \times S_{k_2} \times \cdots \times S_{k_{l_B}})^m \rtimes Z_m$ and $f \in \mathcal{F}_{\frac{\pi}{2}}$ even around $\frac{\pi}{2}$. Then, if m is even, ϕ^* is unstable.*

Proof. We will show the instability of ϕ^* by finding a cut of the network satisfying (2.12). Let $V_0 \subset V(G)$ be the set of nodes within one of the blocks of the

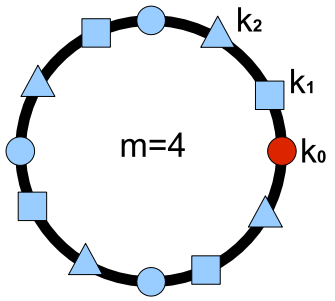


Figure 2.8: Cut of Theorem 2.2, the red block represents one possible set V_0

constellation C_0 and consider the partition induced by V_0, i.e. $P = (V_0, V(G)\backslash V_0)$. Due to the structure of ϕ^*, (2.12) becomes

$$\sum_{ij \in C(P)} f'(\phi_j^* - \phi_i^*) = -k_1 f'(0) + \sum_{l=1}^{l_B} k_l g_m'(\delta_l),$$

where $g_m'(\delta)$ is the derivative of g_m and δ_l is the phase shift between the C_0 and C_l. Finally, since by assumptions $g_m(\delta) \equiv 0 \; \forall \delta$ then it follows that $g_m'(\delta) \equiv 0$ and

$$\sum_{ij \in C(P)} f_{ij}'(\phi_j^* - \phi_i^*) = -k_1 f'(0) < 0.$$

Therefore, by (2.12), ϕ^* is unstable. □

The natural question that arises is whether similar results can be obtained for m odd. The main difficulty in this case is that $g_m(\delta) = 0$ does not hold since we no longer evaluate f at points with phase difference equal to π such that they cancel each other. Therefore, an extra monotonicity condition needs to be added in order to partially answer this question. These conditions and their effects are summarized in the following claims.

Lemma 2.3 (Monotonicity)**.** *Given $f \in \mathcal{F}_{\frac{\pi}{2}} := \mathcal{F}_{\{b=\frac{\pi}{2}\}}$, as in Definition 2.1, such that*

f is strictly concave for $\theta \in [0,\pi]$, then

$$f'(\theta) - f'(\theta - \phi) < 0, \quad 0 \leq \theta - \phi < \theta \leq \pi \tag{2.19}$$

$$f'(\theta) - f'(\theta + \phi) < 0, \quad -\pi \leq \theta < \theta + \phi \leq 0 \tag{2.20}$$

Proof. The proof is a direct consequence of the strict concavity of f. Since $f(\theta)$ is strictly concave then basic convex analysis shows that $f'(\theta)$ is strictly decreasing within $[0,\pi]$. Therefore, the inequality (2.19) follows directly from the fact that $\theta \in [0,\pi]$,$\theta - \phi \in [0,\pi]$ and $\theta - \phi < \theta$. To show (2.20) it is enough to notice that since f is odd ($f \in \mathcal{F}_{\frac{\pi}{2}}$), f is strictly convex in $[\pi, 2\pi]$. The rest of the proof is analogous to (2.19). □

Lemma 2.4 (f' Concavity)**.** *Given $f \in \mathcal{F}_{\frac{\pi}{2}}$ such that f' is strictly concave for $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Then for all $m \geq 4$, $f'(\frac{\pi}{m}) \geq \frac{1}{2}f'(0)$.*

Proof. Since $f'(\theta)$ is concave for $\theta \in [-\pi, \pi]$ then it follows

$$f'(\frac{\pi}{m}) = f'(\lambda_m 0 + (1 - \lambda_m)\frac{\pi}{2}) > \lambda_m f'(0) + (1 - \lambda_m)f'(\frac{\pi}{2}) > \lambda_m f'(0)$$

where $\lambda_m = \frac{m-2}{m}$. Thus, for $m \geq 4$, $\lambda_m \geq \frac{1}{2}$ and

$$f'(\frac{\pi}{m}) > \frac{1}{2}f'(0)$$

as desired. □

Now we show the instability of any equilibria with isotropy $(S_{k_1} \times S_{k_2} \times \cdots \times S_{k_{l_B}})^m \rtimes Z_m$ for m odd and greater or equal to 7.

Theorem 2.3 (Instability for $m \geq 7$ and odd)**.** *Suppose $f \in \mathcal{F}_{\frac{\pi}{2}}$ with f concave in $[0,\pi]$ and f' concave in $[-\frac{\pi}{2}, \frac{\pi}{2}]$, then for all $m = 2k + 1$ with $k \geq 3$ the equilibria ϕ^*s with isotropy $(S_{k_1} \times S_{k_2} \times \cdots \times S_{k_{l_B}})^m \rtimes Z_m$ are unstable.*

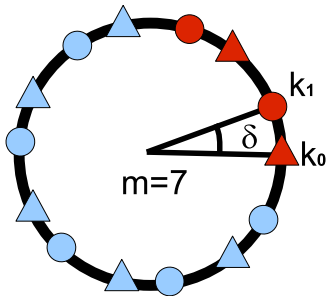


Figure 2.9: Cut used in Theorem 2.3. The dots in red represent all the oscillators of some maximal set S with $d(\phi^*, S) < \frac{4\pi}{m}$

2.3 Effect of Delay

Once delay is introduced to the system of coupled oscillators, the problem becomes fundamentally harder. For example, for pulse-coupled oscillators, the reception of a pulse no longer gives accurate information about the relative phase difference $\Delta\phi_{ij} = \phi_j - \phi_i$ between the two interacting oscillators. Before, at the exact moment when i received a pulse from j, ϕ_j was zero and the phase difference was estimated locally by i as $\Delta\phi_{ij} = -\phi_i$. But now, when i receives the pulse, the difference becomes $\Delta\phi_{ij} = -\phi_i - \psi_{ij}$. Therefore, the delay propagation acts as an error introduced to the phase difference measurement and unless some information is known about this error, it is impossible to predict the behavior. Moreover, as we will see later, slight changes in the distribution can produce nonintuitive behaviors.

Even though it may not be satisfactory for some applications, many existing works choose to ignore delay. (see for e.g., [24, 32, 40]). That is mainly for analytical tractability. On the other hand, when delay is included [118] the studies concentrate on finding bounds on delay that maintain stability.

In this section, we study how delay can *change* the stability in a network of

weakly coupled oscillators. A new framework to study these systems with delay will be set up by constructing an equivalent non-delayed system that has the same behavior as the original one in the continuum limit. We then use this result to show that large heterogeneous delay can help reach synchronization, which is a bit counterintuitive and significantly generalizes previous related studies [25, 136, 137]. We will assume complete graph to simplify notation and exposition although the results can be extended for a broader class of densely connected networks.

The contribution of this section is two fold. Firstly, it improves the understanding of the effect of delays in networks of coupled oscillators. Secondly, it opens new possibilities of using delay based mechanisms to increase the region of attraction of the in-phase equilibrium set. We shall build on existing arguments such as mean field approximation [127] and Lyapunov stability theory [32, 39] while looking at the problem from a different perspective.

2.3.1 Mean Field Approximation

Consider the case of homogeneous oscillators ($\Omega_i = \omega$ and $\omega_i = 0$) with all to all identical coupling ($\mathcal{N}_i = \mathcal{N}\backslash\{i\}, \forall i \in \mathcal{N}$ and $f_{ij} = f\ \forall i, j$). Assume the phase lags ψ_{ij} are randomly and independently chosen from the same distribution with probability density $g(\psi)$. By letting $N \to +\infty$ and $\varepsilon \to 0$ while keeping $\varepsilon N =: \bar{\varepsilon}$ a constant, (2.3) becomes

$$v(\phi, t) := \bar{\varepsilon} \int_{-\pi}^{\pi} \int_{0}^{+\infty} f(\sigma - \phi - \psi) g(\psi) \rho(\sigma, t) d\psi d\sigma, \tag{2.21}$$

where $\rho(\phi, t)$ is a time-variant normalized phase distribution that keeps track of the fraction of oscillators with phase ϕ at time t, and $v(\phi, t)$ is the velocity

field that expresses the net force that the whole population applies to a given oscillator with phase ϕ at time t.

Since the number of oscillators is preserved at any time, the evolution of $\rho(\phi,t)$ is governed by the continuity equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \phi}(\rho v) = 0 \tag{2.22}$$

with the boundary conditions $\rho(0,t) \equiv \rho(2\pi,t)$. Equations (2.21)-(2.22) are not analytically solvable in general. Here we propose a new perspective that is inspired by the following observation.

Consider the non-delayed system of the form

$$\dot{\phi}_i = \varepsilon \sum_{j \in \mathcal{N}_i} H(\phi_j - \phi_i), \tag{2.23}$$

where

$$H(\theta) = f * g(\theta) = \int_0^{+\infty} f(\theta - \psi) g(\psi) d\psi \tag{2.24}$$

is the convolution between f and g.

By the same reasoning of (2.21) it is easy to see that the limiting velocity field of (2.23) is

$$\begin{aligned} v_H(\phi,t) &= \bar{\varepsilon} \int_0^{2\pi} H(\sigma - \phi)\rho(\sigma,t) d\sigma \\ &= \bar{\varepsilon} \int_0^{2\pi} \left(\int_0^{+\infty} f((\sigma - \phi) - \psi) g(\psi) d\psi \right) \rho(\sigma,t) d\sigma \\ &= \bar{\varepsilon} \int_0^{2\pi} \int_0^{+\infty} f(\sigma - \phi - \psi) g(\psi) \rho(\sigma,t) d\psi d\sigma \\ &= v(\phi,t) \end{aligned}$$

where in the first and the third steps we used (2.24) and (2.21) respectively. Therefore, (2.4) and (2.23) **have the same continuum limit.**

Remark 2.4. *Although* (2.23) *is quite different from* (2.4)*, both systems behave exactly the same way in the continuum limit. Therefore, as* N *grows,* (2.23) *approximates* (2.4) *and can be therefore analyzed to understand the behavior of* (2.4)*.*

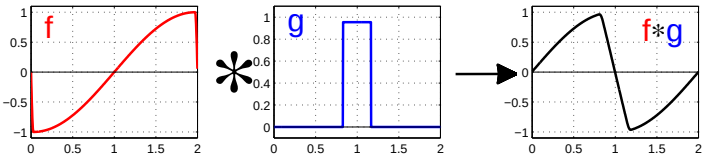


Figure 2.10: Effect of delay in coupling shape

Figure 2.10 shows how the underlying delay (in this case the delay distribution) determines the type of coupling (attractive or repulsive) that produces synchronization. The original function f produces repulsive coupling, whereas the corresponding H is attractive. In fact, as we will soon see, the distribution of delay can not only affect the type of coupling qualitatively, but it can also change the stability of certain phase-locked limit cycles.

We now study two examples to illustrate how this new approximation can provide significant information about performance and stability of the original system. We also provide numerical simulations to verify our predictions.

2.3.2 Kuramoto Oscillators

We start by studying an example in the literature [138] to demonstrate how we can use the previous equivalent non-delayed formulation to provide a better understanding of systems of coupled oscillators with delay. When $f(\theta) = K\sin(\theta)$,

$H(\theta)$ can be easily calculated:

$$\begin{aligned} H(\theta) &= \int_0^{+\infty} K \sin(\theta - \psi) g(\psi) d\psi \\ &= K \int_0^{+\infty} \Im[e^{i(\theta-\psi)} g(\psi)] d\psi = K \Im[e^{i\theta} \int_0^{+\infty} e^{-i\psi} g(\psi) d\psi] \\ &= K \Im\left[e^{i\theta} C e^{-i\xi}\right] = KC \sin(\theta - \xi) \end{aligned}$$

where $\Im$ is the imaginary part of a complex number, i.e. $\Im[a+ib] = b$. The values of $C > 0$ and ξ are calculated using the identity

$$Ce^{i\xi} = \int_0^{+\infty} e^{i\psi} g(\psi) d\psi.$$

This complex number, usually called "order parameter", provides a measure of how the phase-lags are distributed within the unit circle. It can also be interpreted as the center of mass of the lags ψ_{ij}'s when they are thought of as points ($e^{i\psi_{ij}}$) within the unit circle $\mathbb{S}^1$. Thus, when $C \approx 1$, the ψ_{ij}'s are mostly concentrated around ξ. When $C \approx 0$, the delay is distributed such that $\sum_{ij} e^{i\psi_{ij}} \approx 0$.

In this example, (2.23) becomes

$$\dot{\phi}_i = \varepsilon KC \sum_{j \in \mathcal{N}_i} \sin(\phi_j - \phi_i - \xi). \tag{2.25}$$

Here we see how the distribution of $g(\psi)$ has a direct effect on the dynamics. For example, when the delays are heterogeneous enough such that $C \approx 0$, the coupling term disappears and therefore makes synchronization impossible. A complete study of the system under the context of superconducting Josephson arrays was performed [138] for the complete graph topology. There the authors characterized the condition for in-phase synchronization in terms of K and $Ce^{i\xi}$. More precisely, when $KCe^{i\xi}$ is on the right half of the plane ($KC\cos(\xi) > 0$), the system almost always synchronizes. However, when $KCe^{i\xi}$ is on the left half of

the plane ($KC\cos(\xi) < 0$), the system moves towards an incoherent state where all of the oscillators' phases spread around the unit circle such that its order parameter, i.e. $\frac{1}{N}\sum_{l=1}^{N} e^{i\phi_l}$, becomes zero.

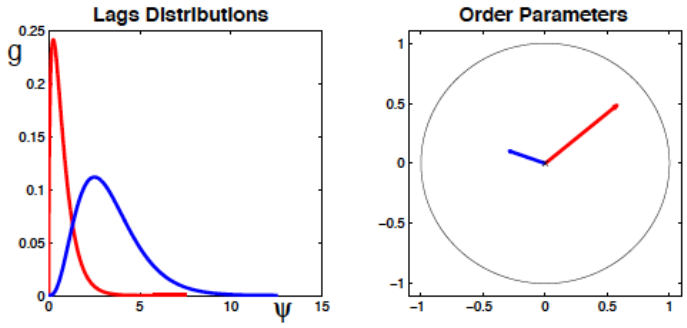


Figure 2.11: Delay distributions and their order parameter $Ce^{i\xi}$

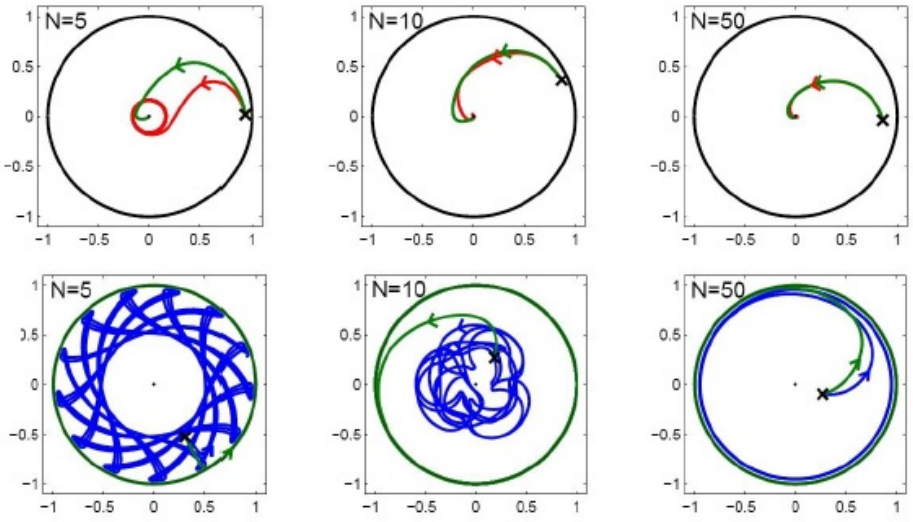


Figure 2.12: Repulsive sine coupling with heterogeneous delays

We now provide simulation results to illustrate how (2.25) becomes a good approximation of the original system when N is large enough. We simulate the original repulsive ($K < 0$) sine-coupled system with heterogeneous delays and its corresponding approximation (2.25). Two different delay distributions, depicted in Figure 2.11, were selected such that their corresponding order parameters lie in different half-planes.

The same simulation is repeated for $N = 5, 10, 50$. Figure 2.12 shows that when N is small, the order parameter of the phases in the original system (in red/blue) draws a trajectory that is completely different with respect to its approximation (in green). However, as N grows, in both cases the trajectories become closer and closer. Since $K < 0$, the trajectory of the system with wider distribution ($C \cos \xi < 0$) drives the order parameter towards the boundary of the circle, i.e., **heterogeneous delay leads to homogeneous phase**.

2.3.3 Effect of Heterogeneity

We now explain a more subtle effect that heterogeneity can produce. Consider the system in (2.23) where H is odd and continuously differentiable. Then, from section 2.2, all the oscillators eventually end up running at the same speed ω with fixed phase difference such that the sum $\sum_{i \in \mathcal{N}_i} H(\phi_j - \phi_i)$ cancels $\forall i$. Moreover, we can apply (2.12) to assess the stability of these orbits. Therefore, if we can find a cut C of the network such that $\sum_{ij \in C} H'(\phi_j^* - \phi_i^*) < 0$, the phase-locked solution will be unstable.

Although this condition is for non-delayed phase-coupled oscillators, the result of this section allows us to translate it for systems with delay. Since H is the convolution of the coupling function f and the delay distribution function g, we can obtain $H'(\phi_j^* - \phi_i^*) < 0$, even when $f'(\phi_j^* - \phi_i^*) > 0$. This usually occurs when the convolution widens the region with a negative slope of H. See Figure 2.10 for an illustration of this phenomenon.

Figures 2.13 and 3.4 show two simulation setups of 45 oscillators **pulse-coupled** all to all. The initial state is close to a phase locked configuration

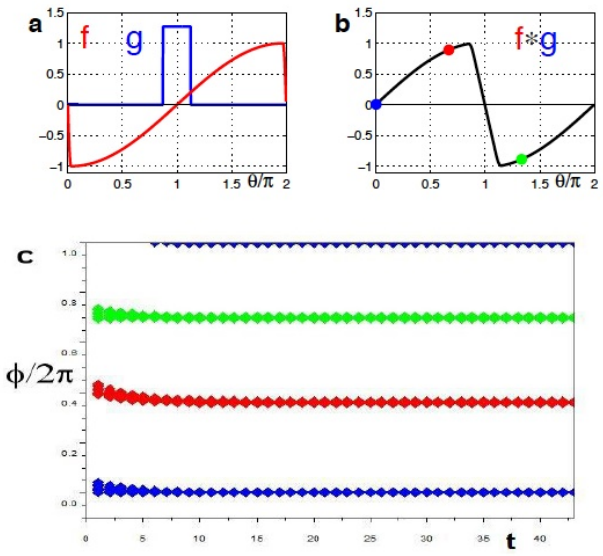


Figure 2.13: Pulse-coupled oscillators with delay: Stable equilibrium

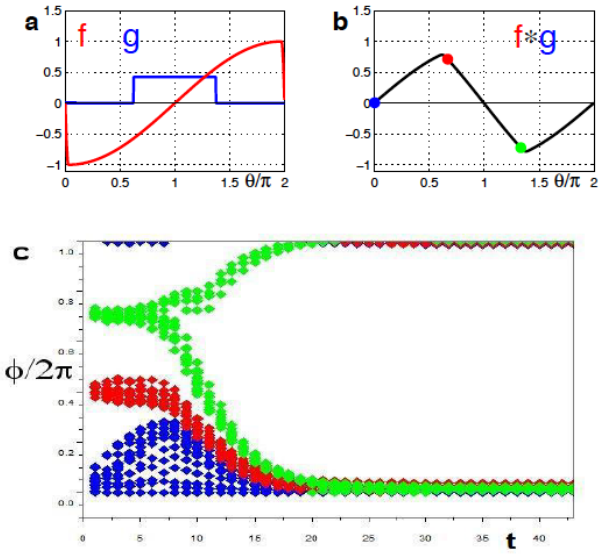


Figure 2.14: Pulse-coupled oscillators with delay: Unstable equilibrium

formed by three equidistant clusters of 15 oscillators each. The shape of the coupling function f and the phase lags distributions are shown in part a. We used (2.5) to implement the corresponding pulse-coupled system (2.1). While f is maintained unchanged between both simulations, the distribution g does change. Thus, the corresponding $H = f * g$ changes as it can be seen in part b; the blue, red, and green dots correspond to the speed change induced in an oscillator within the blue cluster by oscillators of each cluster. Since all clusters have the same number of oscillators, the net effect is zero. In part c the time evolution of oscillators' phases relative to the phase of a blue cluster oscillator are shown. Although the initial conditions are exactly the same, the wider delay distribution on Figure 3.4 produces a negative slope on the red and green points of part b, which destabilizes the clusters and drives oscillators toward in-phase synchrony.

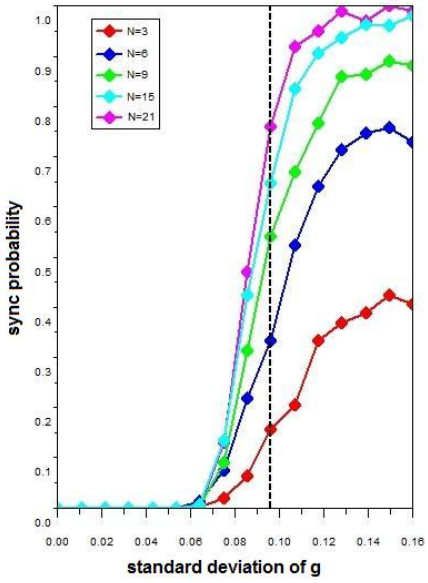


Figure 2.15: Pulse-coupled oscillators with delay: Synchronization probability

Finally, we simulate the same scenario as in Figures 2.13 and 3.4 but now changing N and the standard deviation, i.e. the delay distribution width. Figure 2.15 shows the computation of the synchronization probability vs. standard deviation. The dashed line denotes the minimum value that destabilizes the equivalent system. As N grows, the distribution shape becomes closer to a step, which is the expected shape in the limit. It is quite surprising that as soon as the equilibrium is within the region of H with negative slope, the equilibrium becomes unstable as the theory predicts.

2.4 Heterogeneous Frequencies

We now concentrate on studying the effect of heterogeneous frequencies. As in the previous sections, we are interested in achieving phase consensus. To simplify the analysis we will restrict our attention to continue phase coupled oscillators evolving according to

$$\dot{\phi}_i = \Omega_i + \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j - \phi_i). \tag{2.26}$$

Unfortunately, as soon as the homogeneous frequency assumption is dropped, the problem becomes considerably harder. The challenge is two fold. Firstly, equation (2.4) no longer describes potential dynamics when $\Omega_i \neq 0$ and in fact, to the best of our knowledge, there is not global convergence proof for a system of a finite number of oscillators; the closest result is the work of Ott and Antonsen [139,140] for a system of oscillators in the continuum limit. Secondly, it is not even possible to achieve phase consensus without prior knowledge of

Ω_i. This is mainly due to the fact that in order for synchronization to occur

$$\omega^* = \Omega_i + \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j^* - \phi_i^*), \tag{2.27}$$

must hold $\forall i \in V$ and thus the system needs to compensate the frequency mismatch by introducing a certain phase difference.

Fortunately, if we allow ourselves to modify the dynamics, we can overcome these difficulties by combining ideas from coupled oscillators and linear consensus. Instead of additively changing the frequency as in (2.26), we propose controlling the oscillator speed using a multiplicative scalar γ_i, i.e.

$$\dot{\phi}_i = \Omega_i \gamma_i, \quad \forall i \in V. \tag{2.28}$$

This way, only when $\gamma_i = 1$, the i the oscillator will run at its own natural frequency.

The problem is now reduced to how to define a control law for γ_i. Since our aim is to obtain consensus in both frequency, $\gamma_i \Omega_i$, and phase, ϕ_i, then the adaptation $\dot{\gamma}_i$ should accept such desired solution.

For instance, an initial attempt to solve this problem might be to use

$$\dot{\gamma}_i = \sum_{j \in \mathcal{N}_i} f_{ij}(\phi_j - \phi_i), \quad \forall i \in V,$$

which amounts to adding an integrator to the dynamics. Formally, we can express the dynamics in vector form as,

$$\dot{\gamma} = -BF(B^T \phi) \quad \text{and} \quad \dot{\phi} = \Omega \gamma, \tag{2.29}$$

where $\Omega = \mathrm{diag}[\Omega_i]$.

What is interesting of (2.29) is that even though the frequencies Ω_i might be different, the system still allows phase and frequency consensus. In fact, by

setting $\gamma_i = \frac{\omega^*}{\Omega_i}$, $\phi_i^* = \bar{\phi}$, and integrating (2.29) we obtain the consensus orbit

$$\phi(t) = \omega^* t \mathbf{1}_N + \bar{\phi}\mathbf{1}_N, \forall i \in V.$$

However, a more detailed study of (2.29) unveils an additional oscillatory behavior that this system exhibits. To see this, consider the function $W : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{R}$,

$$W(\phi, \gamma) = \frac{1}{2}\gamma^T \Omega \gamma + V(B^T\phi), \tag{2.30}$$

where $V(y) = \sum_{ij\in E} \int_0^{y_{ij}} f_{ij}(s)ds$.

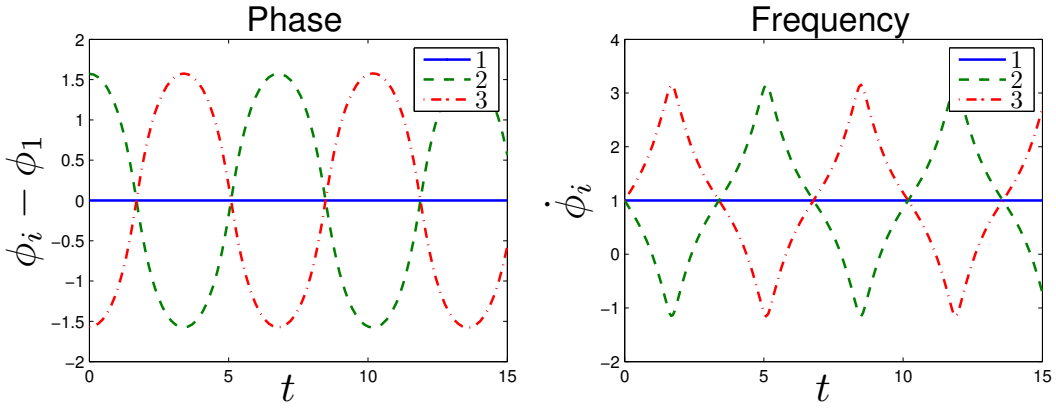
The function $W(\phi, \gamma)$ can be interpreted as the energy function of (2.29). In fact, it is easy to see that $\dot{\phi} = \frac{\partial W}{\partial \gamma}$ and $\dot{\gamma} = -\frac{\partial W}{\partial \phi}$ which means that the system (2.29) is Hamiltonian and that the energy $W(\phi, \gamma)$ remains constant along trajectories, i.e. $\dot{W} \equiv 0$.

This suggests that one could find trajectories in which energy changes from kinetic ($\frac{1}{2}\gamma^T\Omega\gamma$) to potential $V(B^T\phi)$ and back again over time. In Figure 2.16(a) we illustrate one of these trajectories. We simulate a fully connected network of 3 nodes with $\Omega_i = 1, \forall i \in V$ and with initial condition $\phi = (0, \frac{\pi}{2}, -\frac{\pi}{2})^T$, and $\gamma = \mathbf{1}_3^T$.

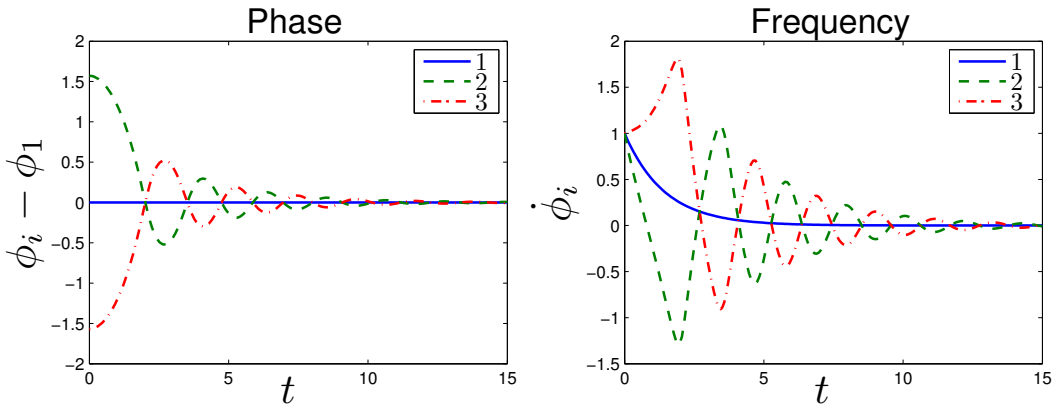
Therefore, although (2.29) allows for the type of solutions we are seeking, the additional integration introduced does not guarantee its convergence. A standard technique to overcome this oscillatory nonlinear behavior [141, 142] is to introduce a damping term in (2.29) that dissipates energy. For instance, consider

$$\dot{\gamma} = -BF(B^T\phi) - \nu\Omega\gamma \quad \text{and} \quad \dot{\phi} = \Omega\gamma, \tag{2.31}$$

where ν is a positive scalar.



(a) Nonlinear oscillations of (2.29): Phases plotted relative to ϕ_1



(b) Adding a positive damping

Figure 2.16: Oscillations and Damping

Figure 2.16(b) shows that the trajectories with the same initial conditions as before now converge. Unfortunately, as Figure 2.16(b) suggests, (2.31) can only admit limit cycles with $\omega^* = 0$ which is unsuitable for our application.

The problem is that the term $-\nu\Omega\gamma$ in $\dot{\gamma}$ is behaving similarly to the system $\dot{x}_i = -\nu x_i$ which clearly has a unique equilibrium in $x_i = 0$. However, if we consider instead,

$$\dot{x}_i = \sum_{ij} a_{ij}(x_j - x_i),$$

it is well known from linear consensus literature that under mild conditions on $a = [a_{ij}]_{ij \in V \times V}$, the trajectories with given initial condition x^0 always converge

to $x_i(t) \to \frac{1}{n}\sum_{i=1}^{n} x_i^0 \; \forall i \in V$. More precisely, this occurs whenever $a_{ij} \geq 0$ and the induced graph $G_a = (V, E_a)$, with $E_a = \{ij \in V \times V | a_{ij} > 0\}$, is *connected*.

Therefore, it seems promising to study

$$\dot{\gamma} = -BF(B^T\phi) - L(a)\Omega\gamma \quad \text{and} \quad \dot{\phi} = \Omega\gamma, \tag{2.32}$$

where $L(a) = B_a \mathrm{diag}[a_{ij}] B_a^T$ is the weighted Laplacian [132] of the possibly different graph $G_a = (V, E_a)$ and B_a denotes the incidence matrix of G_a as defined in (4.1.2).

In the Euclidean counterpart of this problem, it is possible to guarantee convergence even when only two nodes share speed information [143]. In our case, we need to assume that the undirected graph G_a is connected.

Remark 2.5. *One interpretation of the two terms of* $\dot{\gamma}$ *in* (2.32) *is the following. The term* $-BF(B^T\phi)$ *seeks phase consensus, although it cannot achieve it by itself. And the term* $-L(a)\Omega\gamma$ *seeks frequency consensus and it can actually achieve it, but it fails to guarantee phase consensus. Thus, the term* $-L(a)$ *acts as a damping term for the phase consensus algorithm, or equivalently* $-BF(B^T\phi)$ *acts as a correction term of the frequency consensus algorithm. However, there is no simple explanation a priori of why the combination of the two terms can guarantee both frequency and phase consensus simultaneously.*

Remark 2.6. *Another alternative to* (2.32) *is to consider instead*

$$\dot{\gamma} = -BF(B^T\phi) \quad \textit{and} \quad \dot{\phi} = \Omega\gamma - BF(B^T\phi). \tag{2.33}$$

The main advantage of (2.33) *is that it does not need to compute the frequency mismatch which can be very challenging in practice. Studying the convergence properties of this solution is subject of future research.*

2.4.1 Synchronization Frequency

In this section, we compute the value ω^* achieved by (2.32). We start by providing a general characterization for ω^*.

Proposition 2.1. *Given initial conditions* (ϕ^0, γ^0). *If the system* (2.32) *converges to an orbit like* (2.6)*, then the achieved frequency can be computed using*

$$\omega^* = \frac{\sum_{i=1}^N \gamma_i^0}{\sum_{i=1}^N \frac{1}{\Omega_i}}. \tag{2.34}$$

Proof. A well-know property of B (or B_a) is that $\ker[B^T] = \operatorname{span}[\mathbf{1}_N]$ whenever G (or G_a) is connected. Using this property, it is straightforward to show that $\mathbf{1}_N^T \dot{\gamma} \equiv 0$. Then, given initial condition γ^0 we have

$$\begin{aligned}\sum_{i=1}^N \gamma_i(t) = \mathbf{1}_N^T \gamma(t) &= \mathbf{1}_N^T (\gamma^0 + \int_{s=0}^t \dot{\gamma}(s) ds) \\ &= \mathbf{1}_N^T \gamma^0 + 0 = \mathbf{1}_N^T \gamma^0.\end{aligned}$$

Thus, the quantity $\sum_{i=1}^N \gamma_i(t) = \sum_{i=1}^N \gamma_i^0$ is an invariant of the system.

Suppose now that the system converges to a limit cycle, or equivalently that $\gamma_i(t) \to \frac{\omega^*}{\Omega_i}$. Then it follows

$$\sum_{i=1}^N \gamma_i^0 = \sum_{i=1}^N \gamma_i(t) \to \sum_{i=1}^N \frac{\omega^*}{\Omega_i} = \omega^* \sum_{i=1}^N \frac{1}{\Omega_i}.$$

Solving for ω^* gives the desired result. □

When every clock starts with an initial frequency equal to its own natural frequency ($\gamma_i = 1$), ω^* will be the harmonic mean, i.e.,

$$\frac{1}{\omega^*} = \frac{1}{N} \sum_{i=1}^N \frac{1}{\Omega_i}. \tag{2.35}$$

The reason why the system does not achieve the average of $\{\Omega_i\}$ is that the system is in fact averaging a different quantity. This can be seen by substituting Ω_i with $\frac{2\pi}{T_i}$ in (2.35) which gives, $T^* = \frac{2\pi}{\omega^*} = \frac{1}{N}\sum_{i=1}^{N}\frac{2\pi}{\Omega_i} = \frac{1}{N}\sum_{i=1}^{N} T_i$. Thus, the achievable frequency is such that the cycle duration T^* is the average cycle duration among all the oscillators when running with their natural frequencies $\frac{1}{T_i}$'s.

2.4.2 Global Synchronization

We now describe our theoretical convergence results. We will show that under the conditions of Theorem 2.1 the system (2.32) converges to a constant speed orbit with common phase values, i.e. phase and frequency consensus. Throughout this section we will assume that (2.32) contains *isolated orbits*. This is needed in order to guarantee that the system cannot converge to an attractor conformed by a continuum of unstable orbits [144]. Although this property may seem quite restrictive according to example 2.1 and section 2.2.3, if we allow different coupling functions f_{ij}s among different oscillator pairs, the symmetries of the system can be broken and such sets will vanish. Formalizing this idea is the subject of current research.

Frequency Concensus

First, we focus on guaranteeing global convergence towards a constant common frequency, i.e. frequency consensus.

Theorem 2.4 (Frequency Consensus)**.** *Consider the system* (2.32) *running over connected graphs G and G_a, with f_{ij} being symmetric, odd and continuously differentiable.*

Then, for every initial condition, the trajectories converge to a limit cycle as in (2.6) *with* ω^* *as in* (2.34).

Proof. Consider the Lyapunov candidate function $W(\phi, \gamma)$ as defined in (2.30). Notice that the domain of W is composed of the cross product ($\times$) of a compact space $\mathbb{T}^N$ and the unbounded space $\mathbb{R}^N$. Therefore, to apply the global version of Lassale's Invariance Principle, we only need W to be radially unbounded with respect to γ, which is true since Ω is positive definite.

Thus, for any given initial condition (ϕ^0, γ^0) with $W(\phi^0, \gamma^0) = c$ we can always find a scalar $r > 0$ such that for every γ not in a ball $\mathcal{B}_r \subset \mathbb{R}^N$ of radius r and center 0, $W(\phi, \gamma) > c$ for any $\phi \in \mathbb{T}^N$. Therefore, the set $\Psi_c := \{(\phi, \gamma) : W(\phi, \gamma) \leq c\} \subset \mathbb{T}^N \times \mathcal{B}_r$ is compact.

We start by taking the derivative of W along the trajectories. This gives

$$\begin{aligned}
\dot{W}(\phi, \gamma) &= \gamma^T \Omega \dot{\gamma} + \left\langle B \nabla V(B^T \phi), \dot{\phi} \right\rangle \\
&= \gamma^T \Omega [-BF(B^T \phi) - BB^T \Omega \gamma] + \left\langle B \nabla V(B^T \phi), \Omega \gamma \right\rangle \\
&= -\gamma^T \Omega L(a) \Omega \gamma - \gamma^T \Omega B F(B^T \phi) + \gamma^T \Omega B F(B^T \phi) \\
&= -(\Omega \gamma)^T L(a)(\Omega \gamma) \leq 0
\end{aligned}$$

where in the first two steps we use the chain rule for gradients $\nabla(V \circ B^T)(\phi) = B \nabla V(B^T \phi)$ and (2.32), in the third step we use the identity $\nabla V(y) = F(y)$, and in the final step we use the fact that $L(a)$ is positive semidefinite, i.e. $x^T L(a) x \geq 0$ $\forall x$.

Thus, we have shown that Ψ_c is a compact positively invariant set since $\dot{W}(\phi, \gamma) \leq 0$ $\forall (\phi, \gamma) \in \Psi_c$. Lassale's Invariance Principle then implies that the system converges to the largest invariant M set inside $\{\dot{W} \equiv 0\} \cap \Psi_c$. Now, since

G_a connected implies that $\mathbf{1}_N$ is the only eigenvector of $L(a)$ with zero eigenvalue, then $\dot{W} \equiv 0$ implies

$$\Omega\gamma(t) \equiv \omega(t)\mathbf{1}_N.$$

Differentiating both sides, we get $\Omega\dot{\gamma}(t) \equiv \dot{\omega}(t)\mathbf{1}_N$ which is also restricted to $\mathrm{span}[\mathbf{1}_N]$. However, we already know that $\dot{\gamma}(t) \in \ker[\mathbf{1}_N^T]$. Then, since

$$\Omega^{-1}\mathrm{span}[\mathbf{1}_N] \cap \ker[\mathbf{1}_N^T] = \{0\},$$

we must have $\dot{\gamma} \equiv 0$, which implies $\gamma(t) \equiv \omega^*\Omega^{-1}\mathbf{1}_N$ for some constant scalar ω^*. Therefore, we must have $M = M_{\mathbb{T}^N} \times \{\omega^*\Omega^{-1}\mathbf{1}_N\}$ and the system converges to an orbit like (2.6). Proposition 2.1 shows that ω^* is as in (2.34). □

Remark 2.7. *Theorem 2.4 guarantees that the system will synchronize to the harmonic mean of the frequencies (provided $\gamma_i^0 = 1$) but it does not guarantee phase consensus. In other words, Theorem 2.4 places us in the same scenario as remark 2.1. Therefore, it seems reasonable to try to impose a similar condition in the coupling function as the one used in Theorem 2.1.*

Phase Consensus

In this section, we focus on studying the stability of the limit cycles. We know from Theorem 2.4 that (2.32) converges for every initial condition to an orbit like (2.6), where ω^* is characterized by (2.34). Also, since $\gamma(t) \to \gamma^*$ with $\gamma_i^* = \frac{\omega^*}{\Omega_i}$, then from (2.32) we get

$$\begin{aligned} 0 &= -BF(B^T\phi^*) - L(a)\Omega\gamma^* \\ &= -BF(B^T\phi^*) - B_a\mathrm{diag}[a_{ij}]B_a^T\Omega\Omega^{-1}\omega^*\mathbf{1}_N \\ &= -BF(B^T\phi^*) \end{aligned}$$

where in the last step we used again $\ker[B_a^T] = \text{span}[\mathbf{1}_N]$. Thus, ϕ^* must be a solution to $BF(B^T\phi^*) = 0$.

These orbits are exactly the same that would be achieved by the system of coupled oscillators (2.26) if $\Omega_i = \omega^*$ and f_{ij} is as in Theorem 2.4. In section 2.2.2, we showed that their stability when using (2.26) depends on the locations of the eigenvalues of the Laplacian $L(w(\phi^*))$ given in equation (2.13), which is the negation of the Jacobian of (2.26). Thus, if there is at least one negative eigenvalue of $L(w(\phi^*))$, then the orbit defined by ϕ^* is unstable.

Now using Theorem 2.1 we know that whenever $f_{ij} \in \mathcal{F}_b$ and $b \leq \frac{\pi}{N-1}$ every non in-phase orbit will make $L(w(\phi^*))$ have a negative eigenvalue. However, the Jacobian of the new system (2.32) is now

$$J_{\phi^*} = \begin{bmatrix} 0 & \Omega \\ -L(w(\phi^*)) & -L(a)\Omega \end{bmatrix},$$

which now depends on other terms like $L(a)$ and Ω. We will show, however, that provided $L(a)$ is positive definite and induces a connected graph G_a, and $\Omega_i > 0 \; \forall i \in V$, the eigenvalues of $L(w(\phi^*))$ still control the stability.

In order to see this property, consider small perturbation $\delta\phi$, $\delta\gamma$ around a certain orbit (2.6) and the following change of variable

$$x = T^T\delta\phi, \quad z = T^T\Omega\delta\gamma$$

where $T \in \mathbb{R}^{N\times(N-1)}$ is the matrix whose columns $\{T_j\}$ are orthonormal and span $\ker[\mathbf{1}_N^T]$. Notice that by definition, TT^T is the orthogonal projection onto $\ker[\mathbf{1}_N^T]$ and $T^TT = I_{N-1}$, the identity matrix of dimension $N-1$.

The transformation T is clearly not invertible, but it is quite useful to keep track of the disagreement of $\delta\phi$ and $\Omega\delta\gamma$. This is because given $x = T^Tv$, x

becomes zero only when $v \in \mathrm{span}[\mathbf{1}_N]$. .

In other words, the change of variable maps the reference orbit to the point $x = 0$, $z = 0$, and the corresponding dynamics

$$\dot{x} = z \quad \text{and} \quad \dot{z} = T^T \Omega [L(w(\phi^*))Tx + L(a)Tz] \tag{2.36a}$$

describes the evolution of $\delta\phi$ and $\Omega\delta\gamma$ projected onto the subspace $\ker[\mathbf{1}_N^T]$. We now show the following theorem.

Theorem 2.5 (Orbits Instability)**.** *Given connected graphs G and G_a, positive definite Ω and positive semidefinite $L(a)$ with only one zero eigenvalue. Consider any orbit described by ω^* and ϕ^* as in* (2.6)*. Whenever $L(w(\phi^*))$ has a negative eigenvalue, the orbit is unstable and when $L(w(\phi^*))$ is positive semidefinite with one zero eigenvalue, the orbit is stable.*

The proof of this theorem uses a semidefinite version of a Inertia theorem [145, Cor. 2, Th. 5] which is presented here as Lemma 2.5

Lemma 2.5. *Suppose $J \in \mathbb{C}^{p\times p}$ has no eigenvalues on the imaginary axis, $H \in \mathbb{C}^{p\times p}$ is an invertible Hermitian matrix and $JH + HJ^* \geq 0$, where J^* is the conjugate transpose of J. Then the number of eigenvalues of J having positive real part is equal to the number of positive eigenvalues of H.*

Proof. We prove this theorem by showing that if $L(w(\phi^*))$ has a negative eigenvalue, the equilibrium $(x^*, z^*) = (0, 0)$ is unstable. Thus, since x and y are projected version of $\delta\phi$ and $\Omega\delta\gamma$, this shows that in fact the orbit is unstable.

To use Lemma 2.5 we need to show that the Jacobian of (4.7) $\hat{J}$ doesn't have

any eigenvalue on the imaginary axis. Notice first that

$$\hat{J} = \begin{bmatrix} 0 & I_n \\ -\hat{\Omega}\hat{L}_1 & -\hat{\Omega}\hat{L}_2 \end{bmatrix}$$

where $\hat{\Omega} = T^T \Omega T$, $\hat{L}_1 = T^T L(w(\phi^*))T$ and $\hat{L}_2 = T^T L(a)T$. Suppose $\hat{J}$ has an imaginary eigenvalue $j\lambda$, i.e. $\hat{J}v = j\lambda v$ where $v \in \mathbb{C}^{2n}$ and $v = [v_1^T v_2^T]^T$ with $v_1, v_2 \in \mathbb{C}^n$.

$$v_2 = j\lambda v_1$$
$$-\hat{\Omega}\hat{L}_1 v_1 - \hat{\Omega}\hat{L}_2 v_2 = j\lambda v_2$$

If $\lambda = 0$, then $v_1 = v_2 = \mathbf{0}_n$. Otherwise, we obtain

$$\hat{\Omega}\left(\hat{L}_1 + \lambda^2 \hat{\Omega}^{-1} + j\lambda \hat{L}_2\right) v_1 = \mathbf{0}_n \tag{2.37}$$

Now setting $v_1 = x + jy$ and setting real and imaginary part of (2.37) to zero and noticing that $\hat{\Omega}$ is invertible, we obtain

$$Xx - \lambda \hat{L}_2 y = 0 \quad \text{and} \quad Xy + \lambda \hat{L}_2 x = 0$$

with $X = -\hat{L}_1 + \lambda^2 \hat{\Omega}^{-1}$. Therefore, since $\hat{L}_2$ is invertible we obtain $x = \frac{1}{\lambda} X \hat{L}_2^{-1} y$ and follows that

$$\frac{1}{\lambda}\left(\lambda^2 \hat{L}_2 + X \hat{L}_2^{-1} X\right) y = 0$$

which implies that $y = \mathbf{0}_n$ since the facts that $X = X^T$ and L_2 is positive definite make $\lambda^2 \hat{L}_2 + X \hat{L}_2^{-1} X$ positive definite. It follows that $x = v_1 = v_2 = \mathbf{0}_n$. Thus whenever $\hat{J}$ has an imaginary eigenvalue, $v = \mathbf{0}_{2n}$ which is a contradiction.

Finally, consider

$$H = \begin{bmatrix} -\hat{L}_1^{-1} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \hat{\Omega} \end{bmatrix}$$

Then it is easy to show that

$$JH + HJ^T = \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & 2\hat{\Omega}\hat{L}_2\hat{\Omega} \end{bmatrix} \geq 0$$

Result follows then by applying Lemma 2.5. □

Theorem 2.5 provides the connection between our clock synchronization algorithm and equal frequency coupled oscillators. It essentially shows that provided $\Omega_i > 0 \; \forall i$ and $L(a)$ is positive semidefinite with only one zero eigenvalue, both systems contain the same instability condition. This allows us to prove the main result of this section.

Theorem 2.6 (Phase Consensus)**.** *Consider the clock system* (2.32) *running over connected undirected graphs G and G_a. Then, provided $f_{ij} \in \mathcal{F}_b$ with $b \in (0, \frac{\pi}{N-1}]$, for almost every initial condition (ϕ^0, γ^0),* (2.32) *achieves phase and frequency consensus with ω^* as in* (2.34)*.*

Proof. Since G and G_a are connected and f_{ij} by definition is symmetric, odd and continuously differentiable, then by Theorem 2.4, (2.32) will always achieve frequency consensus. As mentioned before, since there are many possible synchronized orbits, this does not guarantee phase consensus.

However, since $f_{ij} \in \mathcal{F}_b$ with $b \in (0, \frac{\pi}{N-1}]$, Corollary 5 of [146] guarantees that any other configuration ϕ^* of (2.6) will produce a negative eigenvalue in $L(w(\phi^*))$. Therefore, by Theorem 2.5, every limit cycle of (2.32) besides the phase consensus one is unstable.

So, unless the initial condition (ϕ, γ) belongs to the zero measure set that converges to these unstable orbits, (2.32) will always converge to the orbit with

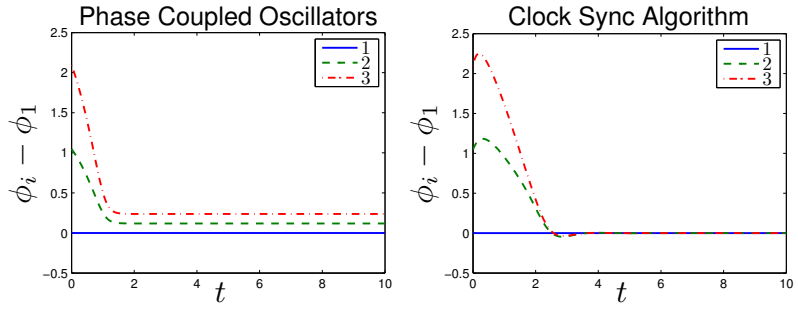
phase and frequency consensus. □

2.4.3 Simulations

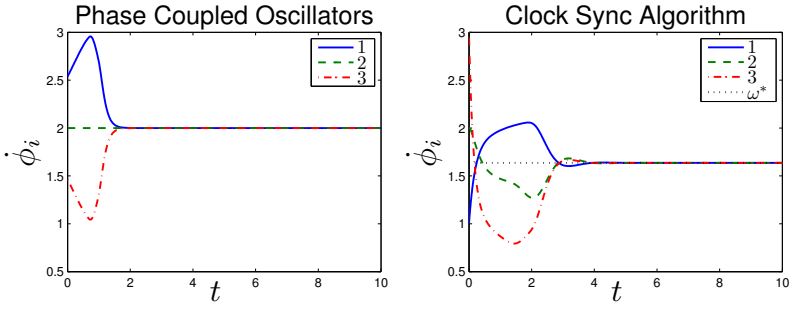
We now present simulations to illustrate our results. In Figure 2.17 we simulate a network of three oscillators running the coupled oscillator algorithm (2.26) and the clock synchronization algorithm (2.32). Both graphs G and G_a are complete and the initial condition is

$$\phi^0 = (0, \frac{\pi}{3}, \frac{2\pi}{3})^T \text{ and } \gamma^0 = (1,1,1)^T,$$

where γ^0 is only used in (2.32). The frequency of each clock is $(\omega_1, \omega_2, \omega_3) = (1,2,3)$.



(a) Phase: Couple Oscillators have to compensate the frequency mismatch

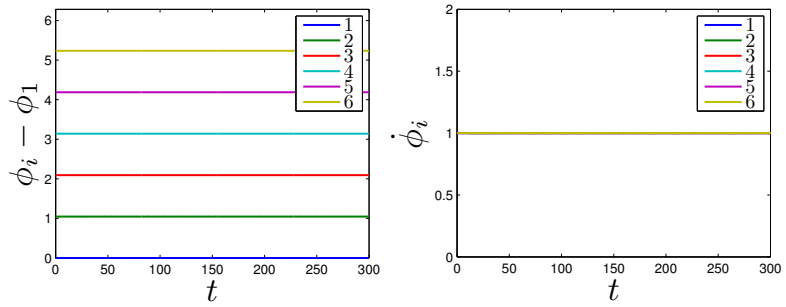


(b) Frequency: Both systems achieve a common frequency

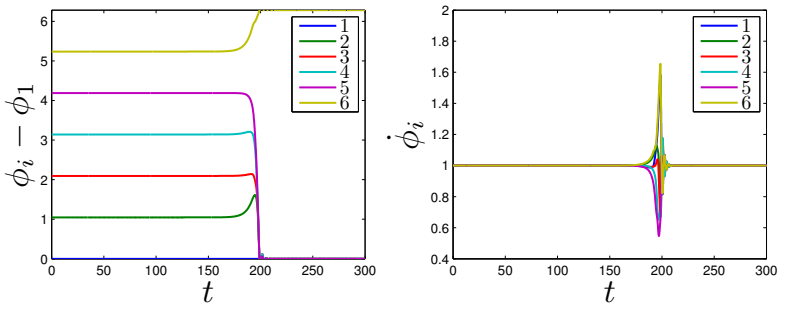
Figure 2.17: Different Frequency CO vs Clock Synchronization

Figure 2.17(a) shows that while (2.32) can achieve phase consensus, (2.26) cannot achieve it due to the frequency difference. Figure 2.17(b) shows that both systems succeed in achieving frequency consensus. Since the initial γ^0 sums to $N = 3$, then (2.32) will have a ω^* as in (2.35), which in our case reduces to $\omega^* = 1.6364$.

We now show why a condition of $b \in (0, \frac{\pi}{N-1}]$ is needed in order to guarantee phase consensus. We simulate (2.32) over a ring network of $N = 6$ nodes, set $\Omega_i = 1\ \forall i \in V$ and initialize the state with values $\phi^0 = [\frac{2\pi k}{6}]_{k\in\{0,\ldots,5\}}$ and $\gamma^0 = \mathbf{1}_6$.



(a) Stable Limit Cycle: $b = \frac{\pi}{2}$



(b) Stable Limit Cycle: $b = \frac{\pi}{6}$

Figure 2.18: De-stablizing orbits by shrinking b below $\frac{\pi}{N-1}$

Figure 2.18 shows two simulations of the same ring network with exactly the same initial conditions. The only difference is the choice of f_{ij}. Figure 2.18(a) shows that when we use a $b = \frac{\pi}{2} > \frac{\pi}{N-1}$ the system stays in the orbit defined by the initial condition. However, once $b = \frac{\pi}{6} < \frac{\pi}{N-1}$, Figure 2.18(b), the orbit is no

longer stable and the system converges to the phase and frequency consensus.

CHAPTER 3

DISTRIBUTED NETWORK CLOCK SYNCHRONIZATION: FUNDAMENTAL LIMITS AND PERFORMANCE OPTIMIZATION

A natural question that arises after developing mechanisms that achieve phase consensus for arbitrary connected graphs as presented in section 2.4 of chapter 2 is whether those results can be applied. Computer clock synchronization seems at first sight an ideal candidate as the main objective in this problem is to make every clock on the network run with the same speed (frequency consensus) and at exactly the same time (phase consensus). However, as in any technology development process, there is a gap between theory and practice. Moreover, very often new challenges appear, while others get simplified.

As we will see in this chapter, the problem of synchronizing computer clocks its not an exception. This chapter is organized as follows. In section 3.1 we provide some background on how clocks are implemented and corrected in computers and how these systems relate with the models of chapter 2. Section 3.2 motivates and describes our algorithm together with an intuitive explanation of why it works. In section 3.3, we analyze the algorithm and determine the set of parameter values and connectivity patterns under which synchronization is guaranteed. The effect of noisy measurement and wander is studied in section 3.4, together with an optimization procedure that finds the parameter values that minimize its effect. Experimental results evaluating the performance of the algorithm are presented in section 3.5.

3.1 Computer Clocks and Synchronization

Most computer architectures keep track of time using a register that is periodically increased by either hardware or kernel's interrupt service routines (ISRs). On Linux platforms, there are usually several different clock devices that can be selected as the clock source by changing the $clocksource$ kernel parameter. One particular counter that has been used recently by several clock synchronization protocols [44,49] is the Time Stamp Counter (TSC), which counts the number of CPU cycles since the last restart. The TSC is a 64-bit counter that has a nominal increment period of δ^o. For example, in the IBM BladeCenter LS21 server, the CPU has a nominal frequency $f^o = 2399.711$MHz which makes $\delta^o = 0.416$ns.

Whenever the counter $c_i(t)$ reaches a value of C it is reset back to zero. Therefore, we can express $c_i(t)$ according to

$$c_i(t) = \lfloor \frac{t - t_i^o}{\delta_i} \rfloor \mod C \tag{3.1}$$

or using the radian units according to

$$\phi_i(t) = 2\pi \frac{c_i(t)}{C} = 2\pi \frac{\lfloor \frac{t-t_i^o}{\delta_i} \rfloor}{C} \mod 2\pi \tag{3.2}$$

where $\lfloor \cdot \rfloor$ is the floor operator, t_i^o is the time when the counter was started and δ_i is the actual CPU cycle period at node i. Without loss of generality, we may assume that all the servers start at the same time, i.e. $t_i^o = t^o$, but they may start with different $c_i(t^o)$ values.

A common assumption is to approximate $c_i(t)$ (and therefore $\phi_i(t)$) by a real variable $c_i(t) = \frac{t-t^o}{\delta_i} \mod C$. This is, in fact, a reasonable approximation since $C = 2^{64} - 1 \gg 1$ and $\delta^o = 0.416\text{ns} \ll 1\mu\text{s}$. Therefore, equations (3.1) and (3.2)

become

$$c_i(t) = \frac{t - t^o}{\delta_i} \mod C \quad \text{and} \quad \phi_i(t) = \Omega_i(t - t^o) \mod 2\pi$$

where the angular frequency $\Omega_i = \frac{2\pi}{\delta_i C}$.

Using this counter, each server can compute its own estimate $x_i(t)$ of the reference time t using

$$x_i(t) = \delta^o c_i(t) + x_i^o = \delta^o \frac{C}{2\pi}\phi_i(t) + x_i^o \tag{3.3}$$

where x_i^o is the estimate of the time when the server was turned on (t_o).

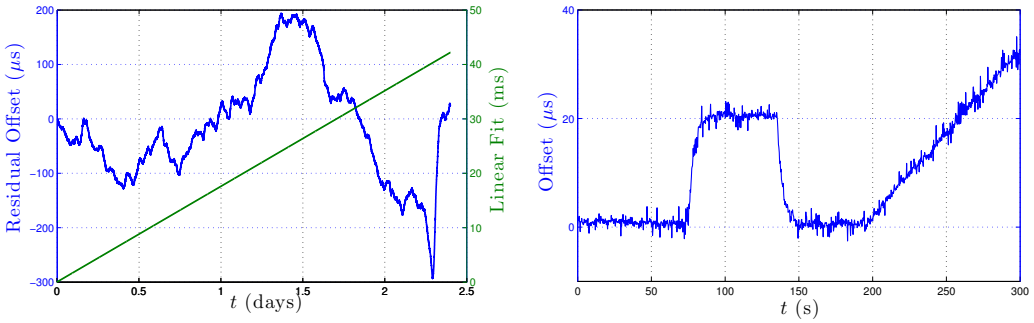
A consequence of (3.3) is that $x_i(t)$ may be interpreted as a properly scaled version of the counter phase $\phi_i(t)$. This could imply, in principle, that in the aim of synchronizing several servers in a network the system may be attracted to a phase-locked state of the form of Figure 2.18(a).

However, this is not a challenge in computer networks due to several reasons. Firstly, even though the counter may reach a value of C, it is fairly easy to modify the routine that handles $x_i(t)$ to identify the counter reset and compensate accordingly by adding $T^o = \delta^o C$ to $x_i(t)$. Secondly, with $C = 2^{64} - 1$ and $\delta^o = 0.416$ns the period of the counter $T^o \approx 243$ years which makes it very unlikely for the counter to reset. Therefore, in this scenario one can assume C to be infinite and disregard any periodic behavior that $c_i(t)$ may incur and transform (3.3) into

$$x_i(t) = r_i(t - t^o) + x_i^o. \tag{3.4}$$

where $r_i := \frac{\delta^o}{\delta_i}$ represents the skew of the local clock with respect to its nominal value; when $r_i > 1$ ($r_i < 1$) the clock is running with frequency higher (lower) than $f^o = \frac{1}{\delta^o}$.

Equation (3.4) is significantly simpler than (3.3). In fact, equation (3.4) is a linear function of the time. Unfortunately, the skew r_i varies due to several factors such as room temperature, mechanical vibrations and interference. This is shown in Figure 3.1(a) where we plot the offset variations between the TSC counters of serv0 and serv1 in our testbed (Figure 3.2, detailed specifics can be found in section 3.5) over more than two days.



(a) Offset between two TSC counters (b) Effect of *adjtimex()* on linux time

Figure 3.1: Comparison between two TSC counters and execution of *adjtimex* command

Therefore, in the same spirit of section 2.4 we will introduce a skew correction s_i in the computation of $x_i(t)$ to compensate the frequency skew. This is equivalent to setting $\dot{\phi}_i = \Omega_i s_i$ as in (2.28) and results in the following expression for $x_i(t)$,

$$x_i(t) = r_i s_i (t - t^o) + x_i^o. \tag{3.5}$$

This map shows explicitly the two fundamental unknowns in a clock synchronization problem (t^o and r_i) and the two parameters that can be used to steer the clock (s_i and x_i^o).

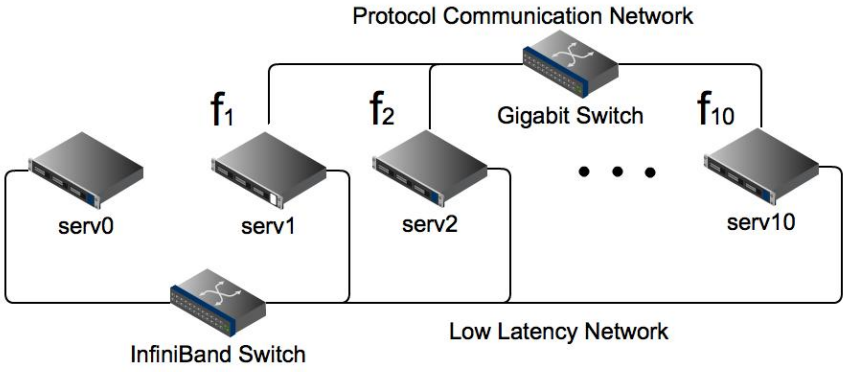


Figure 3.2: Testbed of IBM BladeCenter blade servers

3.1.1 Clock Discipline

To discipline $x_i(t)$ towards t, i.e. make $x_i(t) = t$, one needs to estimate the offset $D_i^x(t) = t - x_i(t)$ at time t^o and the relative frequency error $f_i^{err} = \frac{1-r_i}{r_i}$. In fact, if these values were available at the beginning (something that in practice is not true), then one could just set $s = 1 + f^{err}$ and add an additional offset to (3.5) to get

$$\begin{aligned} x_i(t) &= r_i \left(1 + \frac{1 - r_i}{r_i}\right)(t - t^o) + x_i^o + D_i^x(t^o) \\ &= 1(t - t^o) + x_i^o + (t^o - x_i^o) = t. \end{aligned}$$

Unfortunately, these values are generally unknown and variable. Thus, $f_i^{err}(t)$ and $D_i^x(t)$ need to be repeatedly estimated while the server is running, which introduces several constraints on how the clock can be disciplined as it may affect the execution of time sensitive applications.

To understand the differences between current protocols and relate them to the models of section 2.4, we first rewrite the evolution of $x_i(t)$ based only on

the time instants t_k in which the clock corrections are performed. This marks a clear difference with respect to chapter 2 where the adaptations were done in continuous time. Furthermore, we allow the skew correction s_i to vary over time, i.e. $s_i(t_k)$, and write $x_i(t_{k+1})$ as a function of $x_i(t_k)$. Thus, we obtain

$$x_i(t_{k+1}) = x_i(t_k) + \tau_k r s_i(t_k) + u_i^x(t_k) \tag{3.6}$$

$$s_i(t_{k+1}) = s_i(t_k) + u_i^s(t_k) \tag{3.7}$$

where $\tau_k = t_{k+1} - t_k$ is the time elapsed between adaptations; also known as poll interval [41]. The values $u_i^x(t_k)$ and $u_i^x(t_k)$ represent two different types of correction that a given protocol chooses to do at time t_k and are usually implemented within the interval (t_k, t_{k+1}). $u_i^x(t_k)$ is usually referred to as *offset correction* and $u_i^s(t_k)$ as *skew correction*.

These corrections can be done in Linux OS using the *adjtimex()* interface. The commands

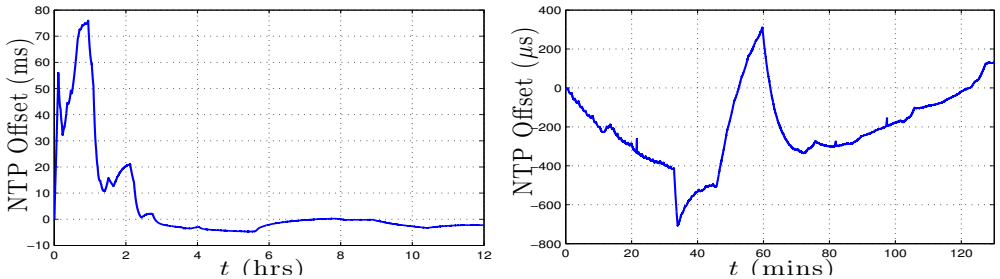
$$adjtimex \text{ -s } offset \text{ and } adjtimex \text{ -f } freq,$$

where *offset* is in nanoseconds (ns) and *freq*= 65536 equals 1ppm (parts per million), are equivalent to setting

$$u^x(t_k) = offset \times 1e^{-9}\text{s and } u^s(t_k) = 1 + \frac{freq}{65536} \times 1e{-}6.$$

Figure 3.1(b) shows the execution of two offset corrections of +20μs and -20μs, and one frequency correction of approx 0.3ppm. We used *offset*= ± 20000 and *freq*= 20000.

Some protocols prefer instead to maintain their own virtual version of $x_i(t)$ as for example IBM CCT [49] and RADclock [44]. This gives more control on how the corrections are implemented since it does not depend on kernel's routines.



(a) NTP initialization period (b) NTP in normal regime

Figure 3.3: Variations of NTP time using TSC as reference

We now proceed to summarize the different types of adaptations implemented by current protocols. The main differences between them are whether they use offset corrections, skew corrections, or both, and whether they update using offset values $D_i^x(t_k) = t_k - x(t_k)$, frequency errors $f_i^{err}(t_k) = \frac{1-r_i s_i(t_k)}{r_i}$, or both.

Offset corrections

This correction consists in using $u_i^s(t_k) = 0$ and either

$$u_i^x(t_k) = \kappa_1 D_i^x(t_k), \text{ or} \tag{3.8}$$

$$u_i^x(t_k) = \kappa_1 D_i^x(t_k) + \kappa_2 f_i^{err}(t_k), \tag{3.9}$$

where $\kappa_1, \kappa_2 > 0$. These adaptations are used by NTPv3 [147] and NTPv4 [41] respectively under ordinary conditions.

The protocols that use (3.8) or (3.9) generally have a slow initialization period, as shown in Figure 3.3(a). This is because the algorithm must first obtain a very accurate estimate of the initial frequency error $f_i^{err}(t^o)$ and set $s_i(t^o) = 1 + f_i^{err}(t^o)$. Furthermore, these updates usually generate non-smooth time evolutions as in Figure 3.3(b) and should be done carefully since they might

introduce backward jumps ($x_i(t_{k+1}) < x_i(t_k)$), which can be problematic for some applications.

Skew corrections

Another alternative that avoids using steep changes in time was proposed in [49]. This alternative does not introduce any offset correction, i.e. $u_i^x(t_k) = 0$, and updates the skew $s_i(t_k)$ using

$$u_i^s(t_k) = \kappa_1 D_i^x(t_k) + \kappa_2 f_i^{err}(t_k). \tag{3.10}$$

In [148] it was shown for a slightly modified version of (3.10) (used $rf^{err}(t_k)$ instead of $f^{err}(t_k)$) that under certain conditions in the parameter values, the algorithm achieves synchronization for very diverse network architectures. Furthermore, equation (3.10) amounts to a discrete time and linear version of (2.32) where we used phase and frequency differences to update s_i.

The main difficulty in using (3.10) is that the estimation of $f^{err}(t_k)$ is nontrivial as it is constantly changing with subsequent updates of $s(t_k)$ and it usually involves sophisticated computations [50,51].

Skew and offset corrections

This type of correction allows dependence on only offset information $D^x(t_k)$ as input to $u^x(t_k)$ and $u^s(t_k)$. For instance, in [45] the update

$$u_i^x(t_k) = \kappa_1 D_i^x(t_k) \quad \text{and} \quad u_i^s(t_k) = \kappa_2 D_i^x(t_k) \tag{3.11}$$

was proposed.

This option allows the system to achieve synchronization without any skew estimation. But the cost of achieving it is introducing offset corrections in $x(t)$, incurring in the same disadvantages discussed in 3.1.1

3.2 Skewless Network Synchronization

We now present an algorithm that overcomes the limitations of the solutions described in Section 3.1. In other words, our solution has the following two properties:

1. Smoothness: The protocol does not introduce steep changes on the time value, i.e. $u^x(t_k) \equiv 0$.
2. Skew independence: The protocol does not use skew information f^{err}as input.

After describing and motivating our algorithm, we show how the updating rule can be implemented in the context of a network environment.

The motivation behind the proposed solution comes from trying to compensate the problem that arises when one tries to naively impose properties 1) and 2), i.e. using

$$u^x(t_k) = 0 \quad \text{and} \quad u^s(t_k) = \kappa D^x(t_k). \tag{3.12}$$

Figure 3.4 shows that this type of clock corrections are unstable; the offset $D^x(t_k)$ of the slave clock oscillates with an exponentially increasing amplitude.

The oscillations in Figure 3.4 arise due to the fundamental limitations of using offset to update frequency. On the other hand, the exponential increase

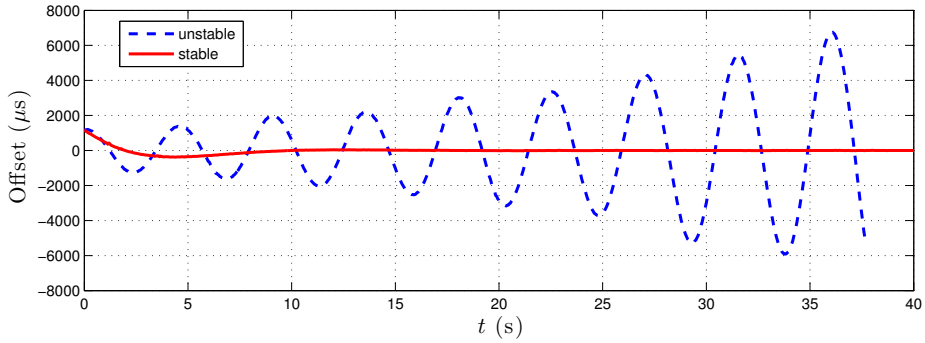


Figure 3.4: Unstable clock steering using only offset information (3.12) and stable clock steering based on exponential average compensation(3.15)

appears since at time t_{k+1} the update is based on the offset at time t_k. Right before updating for the $(k+1)$st time (at t_{k+1}^-), the actual value of the offset has a correction

$$D_i^x(t_{k+1}^-) = D_i^x(t_k) + \tau_k r_i f_i^{err}(t_k)$$

which after noticing that $f_i^{err}(t_k) = f_i^{err}(t_{k+1}^-)$ amounts to an effective correction given by

$$u_i^s(t_k) = \kappa D_i^x(t_k) = \kappa D_i^x(t_{k+1}^-) - \kappa \tau_k r_i f_i^{err}(t_{k+1}^-).$$

Thus, at the moment of the correction, the offset used implicitly includes a negative term in the frequency error that hurts synchronization. This is clearly seen in the case of a slower slave clock $f_i^{err}(t_{k+1}^-) = \frac{1-r_i s_i(t_{k+1}^-)}{r_i} > 0$ with a positive offset $D_i^x(t_{k+1}^-) = t_{k+1}^- - x_i(t_{k+1}^-) > 0$. While the first term of the correction tends to speed up the clock (a desirable effect in this case), the second term tends to slow it down.

One way to try to damp these unstable oscillations is to add a term that opposes the frequency error term. This is done in (3.10) by making $\kappa_2 > \kappa_1 \tau_k r$. However, there are other ways to generate such a term without needing $f^{err}(t_k)$.

For instance, consider the exponentially weighted moving average of the offset

$$y(t_{k+1}) = pD^x(t_k) + (1-p)y(t_k). \tag{3.13}$$

and update $s(t_k)$ using $u^s(t_k) = (\kappa+\gamma)D^x(t_k) - \gamma y(t_k)$.

If we again reference these values at the moment right before the correction (t_{k+1}^-) we have

$$\begin{aligned} u^s(t_k) = \kappa D^x(t_{k+1}^-) - (\kappa+\gamma)\tau_k r f^{err}(t_{k+1}^-) \\ + \gamma(D^x(t_{k+1}^-) - y(t_{k+1}^-)). \end{aligned} \tag{3.14}$$

So now, in the same situation as before ($f^{err}(t_{k+1}^-) > 0$ and $D^x(t_{k+1}^-) > 0$), we have a new term $\gamma(D^x(t_{k+1}^-) - y(t_{k+1}^-))$. This will generally be positive since the offset tends to further increase when the slave clock is slower and thus $D^x(t_{k+1}^-) - y(t_{k+1}^-) > 0$.

Motivated by this discussion, we propose the following updating strategy:

$$u^x(t_k) = 0 \quad \text{and} \quad u^s(t_k) = \kappa_1 D^x(t_k) - \kappa_2 y(t_k) \tag{3.15}$$

where $\kappa_1 = \kappa + \gamma$, $\kappa_2 = \gamma$ and (3.13). Figure 3.4 shows how the proposed strategy can compensate the oscillations without the need to estimate the value of $f^{err}(t_k)$. The stability of the algorithm will depend on how κ_1, κ_2 and p are chosen. A detailed specification of these values is given in Section 3.3.2.

Finally, since we are interested in studying the effect of timing loops, we move away from the client-server configuration implicitly assumed in Section 3.1 and allow mutual or cyclic interactions among nodes. Each node i has its own clock with skew r_i and maintains its own values of $x_i(t_k)$, $s_i(t_k)$ and $y_i(t_k)$. The interactions between different nodes are described by a graph

$G(V,E)$, where V represents the set of nodes ($i \in V$) and E the set of *directed* edges ij; $ij \in E$ means node i can measure its offset with respect to j, $D_{ij}^x(t_k) = x_j(t_k) - x_i(t_k)$.

Within this context, a natural extension of (3.15) is to substitute $D^x(t_k)$ with the weighted average of i's neighbors offsets. Thus, we propose

$$s_i(t_{k+1}) = s_i(t_k) + \kappa_1 \sum_{j \in \mathcal{N}_i} \alpha_{ij} D_{ij}^x(t_k) - \kappa_2 y_i(t_k) \tag{3.16a}$$

$$y_i(t_{k+1}) = p \sum_{j \in \mathcal{N}_i} \alpha_{ij} D_{ij}^x(t_k) + (1-p) y_i(t_k) \tag{3.16b}$$

where $\mathcal{N}_i$ represents the set of neighbors of i and $\alpha_{ij} \neq 0$ iff $j \in \mathcal{N}_i$.

Under this framework, many servers can affect the final frequency of the system. In general, when the system synchronizes globally, we have

$$x_i(t_k) = r^*(t_k - t_0) + x^* \quad i \in V. \tag{3.17}$$

r^* and x^* are possibly different from their ideal values 1 and t_0. Their final values depend on the initial condition of all different clocks as well as the topology, which we assume to be a connected graph in this chapter.

3.3 Convergence Analysis

In this section, we analyze the asymptotic behavior of system (3.16) and provide a necessary and sufficient condition on the parameter values that guarantee its convergence to (3.17). The techniques used are drawn from the control literature, e.g. [45] and [148], yet its application in our case is nontrivial.

Notation 1. *We use $\mathbf{0}_{m\times n}$ ($\mathbf{1}_{m\times n}$) to denote the matrices of all zeros (ones) within $\mathbb{R}^{m\times n}$ and $\mathbf{0}_n$ ($\mathbf{1}_n$) to denote the column vectors of appropriate dimensions. $I_n \in \mathbb{R}^{n\times n}$*

represents the identity matrix. Given a matrix $A \in \mathbb{R}^{n\times n}$ *with Jordan normal form* $A = PJP^{-1}$*, let* $n_A \leq n$ *denote the total number of Jordan blocks* J_l *with* $l \in \mathcal{I}(A) := \{1,...,n_A\}$*. We use* $\mu_l(A)$*,* $l \in \{1,\dots,n\}$ *or just* $\mu(A)$ *to denote the eigenvalues of A, and order them decreasingly* $|\mu_1(A)| \geq \cdots \geq |\mu_n(A)|$*. Finally,* A^T *is the transpose of A,* A_{ij} *is the element of the* i*th row and* j*th column of A,* a_i *is the* i*th element of the column vector a (i.e.* $a = [a_i]^T$*) and* $a^{[i_1,i_2]}$ *is the column sub-vector of elements* a_i*s with* $i_1 \leq i \leq i_2$*.*

It is more convenient for the analysis to use a vector form representation of (3.16) given by

$$z_{k+1} = A_k z_k \tag{3.18}$$

where $z_k := [x(t_k)^T s(t_k)^T y(t_k)^T]^T$ and

$$A_k := \begin{bmatrix} I_n & \tau_k R & 0 \\ -\kappa_1 L & I_n & -\kappa_2 I_n \\ p(-L) & \mathbf{0}_{n\times n} & (1-p)I_n \end{bmatrix},$$

R is the diagonal matrix with elements r_i and L is the Laplacian matrix associated with $G(V,E)$,

$$L_{ii} = \alpha_{ii} := \sum_{j\in\mathcal{N}_i} \alpha_{ij} \text{ and } L_{ij} = \begin{cases} -\alpha_{ij} & \text{if } ij \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The convergence analysis of this section is done in two stages. First, we provide necessary and sufficient conditions for synchronization in terms of the eigenvalues of A_k (Section 3.3.1) and then use Hermite-Biehler Theorem [149] to relate these eigenvalues with the parameter values that can be directly used in practice (Section 3.3.2). All proof details are included in the Appendix for the interested reader.

3.3.1 Asymptotic Behavior

We start by studying the asymptotic behavior of (3.18). That is, we are interested in finding under what conditions the series of $x_i(t_k)$ converge to (3.17).

We will assume without loss of generality that $\tau_k = \tau \; \forall k$ ($A_k = A$) to simplify presentation. The proofs presented here can be readily extended for the time varying τ_k. Thus, we will drop the k index of A_k from here on.

Consider the Jordan normal form [150] of A

$$A = PJP^{-1} := [\zeta_1 \quad \dots \quad \zeta_{3n}] \, J \, [\eta_1 \quad \dots \quad \eta_{3n}]^T \tag{3.19}$$

where $J = \text{blockdiag}(J_l)_{l \in \mathcal{I}(A)}$, ζ_i and η_i are the right and left generalized eigenvectors of A such that

$$\zeta_i^T \eta_j = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

The crux of the analysis comes from understanding the relationship between the multiplicity of the eigenvalue $\mu(A) = 1$ and the eigenvalue $\mu(L) = 0$, and their corresponding eigenvectors.

Lemma 3.1 (Eigenvalues of A and Multiplicity of $\mu(A) = 1$)**.** *A has an eigenvalue $\mu(A) = 1$ with multiplicity 2 if and only if the graph $G(V, E)$ is connected, $\kappa_1 \neq \kappa_2$ and $p > 0$.*

Furthermore, $\mu(A)$ are the roots of

$$g_l(\lambda) := (\lambda - 1)^2(\lambda - 1 + p) + [(\lambda - 1)\kappa_1 + \kappa_2 - \kappa_1]\nu_l \tag{3.20}$$

where $\nu_l = \mu_l(\tau LR)$ and satisfies

$$\nu_n = 0 < |\nu_l| \text{ for } l \in \{1, \dots, n-1\}. \tag{3.21}$$

Lemma 3.2 (Jordan Chains of $\mu(A) = 1$ and $\mu(A) = 1 - p$)**.** *Under the conditions of Lemma 3.1 the right and left Jordan chains, (ζ_1, ζ_2) and (η_2, η_1) respectively, associated with $\mu(A) = 1$ and the eigenvectors ζ_3 and η_3 associated with $\mu(A) = 1 - p$ are given by*

$$[\zeta_1 \ \zeta_2 \ \zeta_3] = \begin{bmatrix} \mathbf{1}_n & \mathbf{1}_n & -\frac{\tau\kappa_2}{p^2}\mathbf{1}_n \\ \mathbf{0}_n & \frac{(R^{-1}\mathbf{1}_n)}{\tau} & \frac{\kappa_2}{p}R^{-1}\mathbf{1}_n \\ \mathbf{0}_n & \mathbf{0}_n & R^{-1}\mathbf{1}_n \end{bmatrix} \text{ and} \tag{3.22}$$

$$[\eta_1 \ \eta_2 \ \eta_3] = \gamma \begin{bmatrix} R^{-1}\xi & \mathbf{0}_n & \mathbf{0}_n \\ -\tau\xi & \xi & \mathbf{0}_n \\ \tau\kappa_2(\frac{1}{p} + \frac{1}{p^2})\xi & -\frac{\kappa_2}{p}\xi & \xi \end{bmatrix} \tag{3.23}$$

where ξ is the unique left eigenvector of $\mu(L) = 0$ and γ is the ξ_i-weighted harmonic mean of r_i, i.e. $\frac{1}{\gamma} = \mathbf{1}_n^T R^{-1}\xi = \sum_{i=1}^n \frac{\xi_i}{r_i}$.

The proof of Lemmas 3.1 and 3.2 can be found in the Appendices A.2 and A.3. We now proceed to state our main convergence result. The proof is relegated to Appendix A.4.

Theorem 3.1 (Convergence)**.** *The algorithm* (3.18) *achieves synchronization for any initial conditions if and only if the graph $G(V, E)$ is connected, $\kappa_1 \neq \kappa_2$, $p > 0$ and $|\mu_l(A)| < 1$ whenever $\mu_l(A) \neq 1$. Moreover, whenever the system synchronizes, we have*

$$x^* = \gamma \sum_{i=1}^n \xi_i \left(\frac{1}{r_i} x_i(t_0) + \tau \frac{\kappa_2}{p^2} y_i(t_0) \right), \text{ and} \tag{3.24a}$$

$$r^* = \gamma \sum_{i=1}^n \xi_i (s_i(t_0) - \frac{\kappa_2}{p} y_i(t_0)). \tag{3.24b}$$

Theorem 3.1 provides an analytical tool to understand the influence of the different nodes of the graph in the final offset x^* and frequency r^*. For example,

suppose that we know that node 1 has perfect knowledge of its own frequency (r_1) and the UTC time at $t = t_0$ ($x_1(t_0) = t_0$), i.e. node 1has perfect knowledge of the UTC time. Then, if we configure the network such that node 1 is the unique leader as the top node in Figures 3.5a and 3.5c, it is easy to show that $\xi_1 = 1$ and $\xi_i = 0 \; \forall i \neq 1$. Using now (3.24a)-(3.24b) and definition of γ we can see that $\gamma = r_1$ and

$$x^* = x_1(t_0) + r_1\tau\frac{\kappa_2}{p^2}y_1(t_0) \text{ and } r^* = r_1 s_1(t_0) - \frac{r_1\kappa_2}{p}y_1(t_0).$$

However, since node 1 knows r_1 and t_0, it can choose $x_1(t_0) = t_0$, $s_1(t_0) = \frac{1}{r_1}$ and $y_1(t_0) = 0$. Thus, we obtain $x^* = t_0$ and $r^* = 1$ which implies by (3.17) that every node in the network will end up with $x_i(t) = t$. In other words, Theorem 3.1 allows us to understand how the information propagates and how we can guarantee that every server will converge to the desired time.

3.3.2 Necessary and sufficient conditions for synchronization

We now provide necessary and sufficient conditions in terms of explicit parameter values (κ_1, κ_2 ,τ and p) for Theorem 3.1 to hold. We will restrict our attention to graphs that have Laplacian matrices with real eigenvalues. This includes for example trees (Figure 3.5a), symmetric graphs with $\alpha_{ij} = \alpha_{ji}$ (Figure 3.5b) and symmetric graphs with a leader (Figure 3.5c).

The proof consists in studying the Schur stability of $g_l(\lambda)$ and it has several steps. We first perform a change of variable that maps the unit circle onto the left half-plane. This transforms the problem of studying the Schur stability into a Hurwitz stability problem which is solved using the Hermite-Biehler Theorem.

Theorem 3.2 (Hurwitz Stability (Hermite-Beihler))**.** *Given the polynomial* $P(s) =$

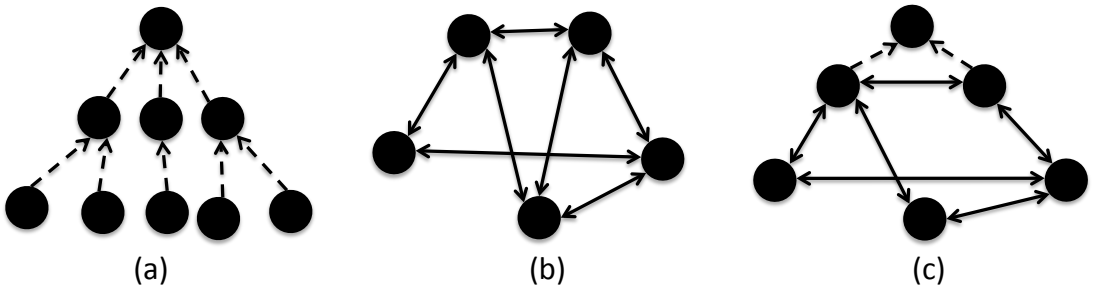


Figure 3.5: Graphs with real eigenvalue Laplacians

$a_n s^n + ... + a_0$, let $P^r(\omega)$ and $P^i(\omega)$ be the real and imaginary part of $P(j\omega)$, i.e. $P(j\omega) = P^r(\omega) + jP^i(\omega)$. Then $P(s)$ is a Hurwitz polynomial if and only if

1. $a_n a_{n-1} > 0$
2. *The zeros of $P^r(\omega)$ and $P^i(\omega)$ are all simple and real and interlace as ω runs from $-\infty$ to $+\infty$.*

Proof. See [149]. □

We now determine the proper parameter values that guarantee synchronization.

Theorem 3.3 (Parameter Values for Synchronization)**.** *Given a connected graph $G(V,E)$ such that the corresponding Laplacian matrix L has real eigenvalues. The algorithm* (3.18) *achieves synchronization if and only if*

(i) $|1 - p| < 1$ *or equivalently* $2 > p > 0$

(ii) $\frac{2\kappa_1}{3p} > \kappa_1 - \kappa_2 > 0$

(iii) $\tau < \frac{p(\kappa_2 - p(\kappa_1 - \kappa_2))}{\mu_{\max}(\kappa_1 - p(\kappa_1 - \kappa_2))^2}$

where $\mu_{\max}$ is the largest eigenvalue of LR.

Even though $\mu_{\max}$ depends on r_i which is in generally unknown, it is easy to show that $\mu_l(LR) \leq \hat{r}_{\max}\mu_l(L)$ where $\hat{r}_{\max}$ is an upper bound of the maximum rate deviation r_i. Furthermore, using Greshgorin's circle theorem, it is easy to show that $\mu_{\max}(L) \leq 2\alpha_{\max} := 2\max_i \alpha_{ii}$. Therefore, if we set

$$\tau < \frac{p(\kappa_2 - \delta\kappa p)}{2\alpha_{\max}\hat{r}_{\max}(\kappa_1 - \delta\kappa p)^2} \tag{3.25}$$

convergence is guaranteed for **every graph** with real eigenvalues.

3.4 Performance Analysis and Optimization

We now focus on studying the performance of our algorithm in the presence of noise. We will consider two possible sources of noise corresponding to measurement errors, due to network congestion, and frequency drifts (*wander*) due to temperature variations, vibrations and interference.

Since our algorithm does not perform skew estimation, the network errors only affect the offset measurements $D_{ij}^x(t_k)$ in (3.16). We use $g_{ij}^w w_{ij}(t_k)$ to denote the error incurred in estimating the offset between nodes i and j at time t_k. This can be produced, for instance, by a congested connection between the two different nodes. We assume that $w_{ij}(t_k)$ has stationary mean $E\left[w_{ij}(t_k)\right] = \bar{w}_{ij}$ $\forall t_k$ and unit variance $E[(w_{ij}(t_k) - \bar{w}_{ij})^2] = 1$ and use g_{ij}^w to weight the different connections.

On the other hand, we model the wander using a time varying rate $r_i(t_k) := r_i + \Delta r_i(t_k)$ where the drift from the mean $\Delta r_i(t_k)$ evolves according to the auto

regressive process

$$\Delta r_i(t_{k+1}) = q_i \Delta r_i(t_k) + g_i^d d_i(t_k) \tag{3.26}$$

where q_i is the autoregression coefficient ($0 < q_i < 1$) and $d_i(t_k)$ is a random variable with zero mean $E[d_i(t_k)] = 0$ and unit variance $E[d_i(t_k)^2] = 1$. Similar models of wander have been used for instance in [51] where $d_i(t_k) \sim \mathcal{N}(0,1)$.

Remark 3.1. *Equation* (3.26) *makes the evolution of* $x_i(t_k)$ *in* (3.18) *nonlinear as now* $x_i(t_{k+1}) = x_i(t_k) + \tau r_i(t_k) s_i(t_k)$. *This is overcome by the fact that* $\Delta r_i(t_k)$ *and* $\Delta s_i(t_k) := s_i(t_k) - s_i^*$ *are of the order of a few parts per millions and therefore* $r_i(t_{k+1}) s_i(t_{k+1})$ *is approximated by*

$$\begin{aligned} r_i(t_{k+1}) s_i(t_{k+1}) - r_i s_i^* &\approx \Delta r_i(t_{k+1}) s_i^* + r_i \Delta s_i(t_{k+1}) \\ &= s_i^*(q_i \Delta r_i(t_k)) + r_i(\Delta s_i(t_k) + u_i^s(t_k) + \beta_i g_i^d d_i(t_k)) \end{aligned} \tag{3.27}$$

where $\beta_i = \frac{s_i^*}{r_i} \approx 1$. *Equation* (3.27) *also shows that we can equivalently assume that* $d_i(t_k)$ *is a noise source that affects* $s_i(t_{k+1})$ *instead of* $r_i(t_{k+1})$.

This motivates the study of the stochastic process

$$z_{k+1} = A z_k + B e_k \tag{3.28a}$$

$$v_{k+1} = C z_k \tag{3.28b}$$

where $e_k = [w_k^T \; d_k^T]^T$, $B = [B_w \; B_d]$ with

$$B_w = \begin{bmatrix} \mathbf{0}_{n\times m} \\ -\kappa_1 B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w] \\ -p B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w] \end{bmatrix}, \quad B_d = \begin{bmatrix} \mathbf{0}_{n\times n} \\ \mathrm{diag}[\beta_i g_i^d] \\ \mathbf{0}_{n\times n} \end{bmatrix},$$

$B_G^- = \min\{B_G, \mathbf{0}_{n\times m}\}$ and B_G being the incidence matrix of $G(V,E)$ [1] and $w_k = [w_{ij}(t_k)]^T$. The matrix C maps the system state z_k to the performance metric v_k and will be specified in Section 3.4.

[1] Notice that using this definition $L = B_G^- \mathrm{diag}[\alpha_{ij}] B_G^T$

In the remainder of this section, we first study the effect of biased network noise ($\bar{w}_{ij} \neq 0$) in the asymptotic frequency of the system and time offsets. In particular, we show that for arbitrarily distributed noise with stationary mean, the system's frequency tends to constantly drift unless there is a well defined leader in the topology. We then proceed to study how the parameters and network topology affect system performance, which is represented by the output signal v_k of the stochastic process.

We will assume that the input is white noise, i.e. $E[e_k e_l^T] = I_{m+n}\delta(l-k)$,[2] and focus on reducing the output power $||v_k||_2^2 = \lim_{N\to+\infty} \frac{1}{N} \sum_{k=0}^{N-1} v_k^T v_k$. This is known in the control theory community as $\mathcal{H}_2$ optimal control.

3.4.1 Frequency Drift and Time Offset

We now concentrate on studying the evolution of the first moment of the stochastic process (3.28). That is, we want to understand how $\bar{z}_k = E[z_k]$ evolves as $k \to +\infty$. To simplify the analysis, consider the change of variable $\hat{z} = P^{-1}\bar{z} = [\eta_1 \quad \dots \quad \eta_{3n}]^T \bar{z}$ where P is defined as in (3.19). This change of variable further simplifies the dynamics of (3.28) giving

$$\hat{z}_{k+1} = J\hat{z}_k + P^{-1}B_w\bar{w}. \tag{3.29}$$

Notice that we assume $\bar{d} = \mathbf{0}_n$ and thus the term $B_d\bar{d}$ is omitted from (3.29).

While it is difficult to provide a physical interpretation of most of the variables of the vector $\hat{z}$, it is possible to relate certain groups of states to different roles within the system. Consider the following partition of the state space

[2] $\delta(k) = 1$ if $k = 0$ and 0 o.w.

$\hat{z} = [(\hat{z}^{[1,3]})^T | (\hat{z}^{[4,3n]})^T]^T$. By definition of $\hat{z}$ and (A.6) we have

$$\hat{z}_{k+1}^{[1,3]} = \hat{J}_1 \hat{z}_k^{[1,3]} + [\eta_1 \ \eta_2 \ \eta_3]^T B_w \bar{w} \tag{3.30}$$

$$\hat{z}_{k+1}^{[4,3n]} = \hat{J}_2 \hat{z}_k^{[4,3n]} + [\eta_4 \ \dots \ \eta_{3n}]^T B_w \bar{w} \tag{3.31}$$

where

$$\hat{J}_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1-p \end{bmatrix} \text{ and } \rho(\hat{J}_2) < 1.$$

The function $\rho(A)$ is the spectral radius of A or equivalently the largest absolute value of its eigenvalues.

The following lemma is crucial in understanding the role of the different states of $\hat{z}$. Let

$$\delta\bar{x}_k := x(t_k) - \mathbf{1}_n \left(\hat{z}_1(t_k) + \hat{z}_2(t_k) - \frac{\tau\kappa_2}{p^2} \hat{z}_3(t_k) \right) \tag{3.32}$$

$$\delta\bar{s}_k := s(t_k) - R^{-1}\mathbf{1}_n \left(\frac{1}{\tau} \hat{z}_2(t_k) + \frac{\kappa_2}{p} \hat{z}_3(t_k) \right) \tag{3.33}$$

$$\delta\bar{y}_k := y(t_k) - R^{-1}\mathbf{1}\hat{z}_3(t_k). \tag{3.34}$$

Lemma 3.3 (Mean Convergence)**.** *Under the conditions of Theorem 3.1 the system* (3.28) *converges in mean towards*

$$\delta\bar{x}_k \to \delta\bar{x}^*, \quad \delta\bar{s}_k \to \delta\bar{s}^* \text{ and } \delta\bar{y}_k \to \delta\bar{y}^* \tag{3.35}$$

with

$$\delta\bar{z}^* = \left[\delta\bar{x}^{*T} \ \delta\bar{s}^{*T} \ \delta\bar{y}^{*T} \right]^T = [\zeta_4 \ \dots \ \zeta_{3n}] \ \hat{z}^{[4,3n]*} \tag{3.36}$$

and

$$\hat{z}^{[4,3n]*} = (I - \hat{J}_2)^{-1} [\eta_4 \ \dots \ \eta_{3n}]^T B_w \bar{w}. \tag{3.37}$$

Proof. Since $\hat{J}_2$ has $\rho(\hat{J}_2) < 1$ then $z^{[4,3n]}(t_k)$ converges for every initial condition to a unique value which is the fix point of (3.31) given by $\hat{z}^{[4,3n]*}$(3.37) or in terms of the original system variables by $\delta\bar{z}^*$ (3.36).

Now by definition of $\hat{z}$,

$$\bar{z}_k = P\hat{z}_k = \sum_{l=1}^{3} \zeta_l \hat{z}_l(t_k) + [\zeta_4 \ \dots \ \zeta_{3n}]\, \hat{z}_k^{[4,3n]}$$

Then, since $\hat{z}_k^{[4,3n]} \to \hat{z}_k^{[4,3n]*}$ we have

$$\bar{z}_k - \sum_{l=1}^{3} \zeta_l \hat{z}_l(t_k) \to [\zeta_4 \ \dots \ \zeta_{3n}]\, \hat{z}^{[4,3n]*} = \delta\bar{z}^*. \tag{3.38}$$

Thus, by Lemma 3.2 we obtain

$$\begin{bmatrix} \bar{x}_k - \delta\bar{x}^* \\ \bar{s}_k - \delta\bar{s}^* \\ \bar{y}_k - \delta\bar{y}^* \end{bmatrix} - \begin{bmatrix} \mathbf{1}_n(\hat{z}_1(t_k) + \hat{z}_2(t_k) - \frac{\tau\kappa_2}{p^2}\hat{z}_3(t_k)) \\ R^{-1}\mathbf{1}_n\left(\frac{1}{\tau}\hat{z}_2(t_k) + \frac{\kappa_2}{p}\hat{z}_3(t_k)\right) \\ R^{-1}\mathbf{1}_n\hat{z}_3(t_k) \end{bmatrix} \to \mathbf{0}_{3n}$$

which is equivalent to (3.35). □

Lemma 3.3 shows that while $\hat{z}^{[1,3]}$ has an homogeneous and (possibly) non-constant effect on every node, $\hat{z}^{[4,3n]}$ in the limit introduces a fixed offset. In particular, when $\bar{w} = 0$ (e.g. zero mean noise) $\delta\bar{z}^* = 0$, $\hat{z}_2(t_k) = (\hat{z}_2)_0$, $\hat{z}_1(t_k) = (\hat{z}_1)_0 + k(\hat{z}_2)_0$ and $\hat{z}_3(t_k) \to 0$ achieving time consensus as in Theorem 3.1.

The next two theorems summarize the main results of this section.

Theorem 3.4 (Frequency Drift)**.** *In the presence of noise and under the condition of Theorem 3.1 the system synchronizes in mean with constant frequency if and only if*

$$\sum_{i=1}^{n} \xi_i \sum_{j\in\mathcal{N}_i} \alpha_{ij} g_{ij}^{w} \bar{w}_{ij} = 0. \tag{3.39}$$

Moreover, when this happens the mean frequency r^* is given by (3.24b)*.*

Proof. By Lemma 3.3 we know that $\bar{s}_i(t_k)$ asymptotically approaches $\delta\bar{s}_i^* + \frac{1}{\tau r_i}\hat{z}_2(t_k) + \frac{\kappa_2}{pr_i}\hat{z}_3(t_k)\ \forall i$. Therefore, $\bar{s}_i(t_k)$ becomes constant if and only if $\frac{1}{\tau r_i}\hat{z}_2(t_k) + \frac{\kappa_2}{p}\hat{z}_3(t_k)$ does.

Now from (3.30) it follows that

$$\begin{aligned}
\hat{z}_2(t_{k+1}) &= \hat{z}_2(t_k) + \eta_2^T B_w \bar{w} \\
&= \hat{z}_2(t_k) - \tau\gamma(\kappa_1 - \kappa_2)\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} \\
\hat{z}_3(t_{k+1}) &= (1-p)\hat{z}_3(t_k) + \eta_3^T B_w \bar{w} \\
&= (1-p)\hat{z}_3(t_k) - p\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w}.
\end{aligned}$$

Thus, $z_3(t_k) \to -\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w}$ and

$$\hat{z}_2(t_k) = \hat{z}_2(0) + t_k\gamma(\kappa_2 - \kappa_1)\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w}$$

which is constant if and only if $(\kappa_2 - \kappa_1)\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} = 0$. But since, by Theorem 3.1, $\kappa_1 > \kappa_2$ then we must have

$$0 = -\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} = \sum_{i=1}^{n} \xi_i \sum_{j \in \mathcal{N}_i} \alpha_{ij} g_{ij}^w \bar{w}_{ij}.$$

□

Notice that (3.39) implies that $\hat{z}_k^{[1,3]}$ behaves identically to the noiseless version

It is important to highlight the relationship between (3.39) and the topology of G. In particular, it is possible to differentiate two different scenarios in which (3.39) can be satisfied.

1. G has a unique leader (say $i = 1$): In this case we have $\mathcal{N}_1 = \emptyset$, i.e. $\alpha_{1j} = 0$ $\forall j$, $\xi_1 = 1$ and $\xi_j = 0\ \forall j \neq 1$. That is $-\xi^T B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} = \xi_1 0 = 0$

2. G does not have a well defined root: Thus, there are at least two nodes with $\xi_i \neq 0$ and $\bar{w}$ is such that $\xi^T B_G^- \text{diag}[\alpha_{ij} g_{ij}^w] \bar{w} = 0$.

Thus, while condition 1) can be satisfied by a proper configuration of the network, condition 2) is only satisfied by a set of values of $\bar{w}$ with zero measure. Therefore, in practice the only possible way to avoid frequency drift is by using a graph G with a well defined leader.

Corollary 3.1 (Frequency Robustness)**.** *In the presence of measurements noise, the mean frequency $r_i \bar{s}_i(t_k)$ converges to a fixed value* (3.24b) *with probability one (in the set of possible $\bar{w}$) if and only if G has a well defined leader.*

Furthermore, while at first sight it seems difficult to evaluate $\delta\bar{z}^*$ using (3.37), the following Theorem provides us with a physical interpretation.

Theorem 3.5 (Time Offsets)**.** *Under the conditions of Theorem 3.1 and* (3.39)*, $\delta\bar{z}^*$ in* (3.36) *becomes*

$$\delta\bar{z}^* = \begin{bmatrix} -N_1 L^\dagger B_G^- \text{diag}[\alpha_{ij}] \bar{w} \\ \mathbf{0}_n \\ \mathbf{0}_n \end{bmatrix}$$

where $L^\dagger$ is the pseudo inverse of L and $N_1 = (I_n - \gamma \mathbf{1} \xi^T R^{-1})$.

Proof. By Lemma 3.2 and definition of $\hat{z}$ we can compute

$$\begin{aligned} \delta\bar{x}_k &= x(t_k) - \mathbf{1}_n \left(\hat{z}_1(t_k) + \hat{z}_2(t_k) - \frac{\tau\kappa_2}{p^2} \hat{z}_3(t_k) \right) = x(t_k) \\ &-\gamma \mathbf{1}_n \left(\xi^T R^{-1} x(t_k) - \tau \xi^T s(t_k) + \tau\kappa_2 \left(\frac{1}{p} + \frac{1}{p^2} \right) \xi^T y(t_k) \right) \\ &-\gamma \mathbf{1}_n \left(\tau \xi^T s(t_k) - \frac{\tau\kappa_2}{p} \xi^T y(t_k) \right) + \gamma \mathbf{1}_n \frac{\tau\kappa_2}{p^2} \xi^T y(t_k) \\ &= x(t_k) - \gamma \mathbf{1}_n \xi^T R^{-1} x(t_k) = N_1 x(t_k). \end{aligned}$$

Similarly, we have $\delta\bar{s}_k = N_2 s(t_k)$ and $\delta\bar{y}_k = N_2 y(t_k)$ where $N_2 = (I_n - \gamma R^{-1}\mathbf{1}\xi^T)$.

Moreover, since $N_1 R = R N_2$, $N_1 L = L N_2 = L$ and $N_2 B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} = B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w}$ (by (3.39)) we have

$$\delta\bar{x}_{k+1} = \delta\bar{x}_k + \tau R \delta\bar{s}_k \tag{3.40a}$$

$$\delta\bar{s}_{k+1} = -\kappa_1 L\delta\bar{x}_k + \delta\bar{s}_k - \kappa_2\delta\bar{y}_k - \kappa_1 B_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} \tag{3.40b}$$

$$\delta\bar{y}_{k+1} = -pL\delta\bar{x}_k + (1-p)\delta\bar{y}_k - pB_G^- \mathrm{diag}[\alpha_{ij} g_{ij}^w]\bar{w} \tag{3.40c}$$

Now, by Lemma 3.3 we know that (3.35) holds and therefore $\delta\bar{z}^*$ is a fixed point of (3.40). Thus, (3.40a) implies that $\delta\bar{s}^* = 0$ and (3.40b)$-\frac{\kappa_1}{p}$(3.40c) gives

$$(\kappa_1 - \kappa_2)\delta\bar{y}^* = 0$$

which implies $\delta\bar{y}^* = 0$ since $\kappa_1 > \kappa_2$. Finally using (3.40c) again we have

$$\begin{aligned} L\delta\bar{x}^* + B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} &= 0 \\ L^\dagger L\delta\bar{x}^* &= -L^\dagger B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} \\ N_3\delta\bar{x}^* &= -L^\dagger B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} \end{aligned}$$

where $N_3 = L^\dagger L = (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T)$.

Thus, since $N_1 N_3 = N_1$ and by definition $N_1\delta\bar{x} = N_1^2\bar{x} = N_1\bar{x} = \delta\bar{x}$ it follows that

$$\begin{aligned} N_3\delta\bar{x}^* &= -L^\dagger B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} \\ N_1 N_3\delta\bar{x}^* &= -N_1 L^\dagger B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} \\ \delta\bar{x}^* &= -N_1 L^\dagger B_G^- \mathrm{diag}[\alpha_{ij}]\bar{w} \end{aligned}$$

□

Remark 3.2 (Convergent Measures)**.** *One interesting property of the system* (3.18) *that is evidenced by theorem 3.5 is that even in the absence of noise* $\bar{w} = 0$ *the system appears to converge each individual state* $x_i(t_k)$*,* $s_i(t_k)$ *and* $y_i(t_k)$ *toward the scalar values* $\tilde{x}(t_k)$*,* $\frac{1}{r_i}\tilde{s}(t_k)$ *and* $\frac{1}{r_i}\tilde{y}(t_k)$ *where*

$$\tilde{x}(t_k) := \gamma \xi^T R^{-1} \bar{x}(t_k), \quad \tilde{s}(t_k) := \gamma \xi^T \bar{s}(t_k), \quad \tilde{y}(t_k) := \gamma \xi^T \bar{y}(t_k)$$

and evolve according to

$$\begin{bmatrix} \tilde{x}(t_{k+1}) \\ \tilde{s}(t_{k+1}) \\ \tilde{y}(t_{k+1}) \end{bmatrix} = \begin{bmatrix} 1 & \tau & 0 \\ 0 & 1 & -\kappa_2 \\ 0 & 0 & (1-p) \end{bmatrix} \begin{bmatrix} \tilde{x}(t_k) \\ \tilde{s}(t_k) \\ \tilde{y}(t_k) \end{bmatrix}$$

Therefore, it is possible to interpret $\tilde{x}(t_k)$*,* $\tilde{s}(t_k)$ *and* $\tilde{y}(t_k)$ *as a generalization of the invariant measure used in section 2.4 to compute the synchronizing frequency* ω^**. Here, instead of a constant (invariant) measure, we have three convergent measures that define the global behavior of the system.*

3.4.2 $\mathcal{H}_2$ Performance Optimization

We now proceed to study the effect of noisy measurements and wander on the output standard deviation of the system ($||v_k||_2$) when the input e_k is white noise ($E[e_k e_l^T] = I_{m+n}\delta(l-k)$). In other words, we seek to minimize

$$f(\kappa_1, \kappa_2, p, \alpha_{ij}) = ||v_k||_2 = \sqrt{E\left[\lim_{N\to+\infty} \frac{1}{N} \sum_{k=0}^{N-1} v_k^T v_k\right]}$$

Since in practice we want to avoid any frequency drift introduced by the noise, we will assume in this section that (3.39) holds. Thus, all the randomness

of the system is concentrated in $\delta x_k = N_1 x(t_k)$, $\delta s_k = N_2 s(t_k)$ and $\delta y_k = N_2 y(t_k)$ and we only study the stochastic process

$$\delta z_{k+1} = NA\delta z + NBe_k$$

$$v_{k+1} = C\delta z_k$$

where $N = \text{blockdiag}(N_1, N_2, N_2)$.

This optimization problem is standard in the control theory community and it can be shown to be equivalent to

$$\min_{X,\kappa_1,\kappa_2,p,\alpha_{ij}} \quad f(\kappa_1,\kappa_2,p,\alpha_{ij}) := \sqrt{\text{trace}[XBNN^TB^T]} \tag{3.41a}$$

$$\text{subject to} \quad \rho(NA) \leq \rho^* \tag{3.41b}$$

$$X = A^T N^T X N A + C^T C \tag{3.41c}$$

where A is a function of $(\kappa_1, \kappa_2, p, \alpha_{ij})$ and $\rho^* < 1$. The constraint (3.41b) has been added in order to maintain the stability of A.

While it is not generally easy to find the global minimum of (3.41), there has been intensive research to study the continuous time [151] and discrete time [152] versions of the optimization problem

$$\min_{K,X} \quad f(K) := \sqrt{\text{trace}[X\bar{B}\bar{B}^T]} \tag{3.42a}$$

$$\text{subject to} \quad \rho(\bar{A}) \leq \rho^* \tag{3.42b}$$

$$X = \bar{A}^T X \bar{A} + \bar{C}^T \bar{C} \tag{3.42c}$$

where $\bar{A} := \hat{A} + \hat{B}_2 K \hat{C}_2$, $\bar{B} := \hat{B}_1 + \hat{B}_2 K \hat{D}_{21}$ and $\bar{C} := \hat{C}_1$ and δz_k is interpreted as evolving according to the closed loop standard form system

$$\delta z_{k+1} = (\hat{A} + \hat{B}_2 K \hat{C}_2)\delta z_k + (\hat{B}_1 + \hat{B}_2 K \hat{D}_{21})e_k$$

$$v_k = \hat{C}_1 \delta z_k,$$

with K being the static output feedback matrix.

Proposition 3.1. *The optimization problem* (3.41) *can be written as* (3.42) *with*

$$\hat{A} = N,\ \hat{C}_1 = C, \hat{C}_2 = \begin{bmatrix} B_G^T & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times n} & I_n & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & I_n \end{bmatrix},$$

$$\hat{B}_2 = \begin{bmatrix} N_1 R & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & B_G^- & N_2 & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} & B_G^- & N_2 \end{bmatrix},$$

$$\hat{B}_1 = \begin{bmatrix} \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times m} & \mathrm{diag}[g_i^d] \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \end{bmatrix},\ \hat{D}_{21} = \begin{bmatrix} \mathrm{diag}[g_{ij}^w] & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \end{bmatrix},$$

$$\textit{and } K = \begin{bmatrix} \mathbf{0}_{n\times m} & \tau I_n & \mathbf{0}_{n\times n} \\ -\kappa_1 \mathrm{diag}[\alpha_{ij}] & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} & -\kappa_2 I_n \\ -p\mathrm{diag}[\alpha_{ij}] & \mathbf{0}_{m\times n} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} & -pI_n \end{bmatrix}$$

Proof. The proof of this proposition is simple computation. By definition of $\hat{B}_2$, K and $\hat{C}_2$

$$\hat{B}_2 K C_2 = \begin{bmatrix} \mathbf{0}_{n\times m} & \tau N_1 R & \mathbf{0}_{n\times n} \\ -\kappa_1 B_G^- \mathrm{diag}[\alpha_{ij}] B_G^T & \mathbf{0}_{n\times n} & -\kappa_2 N_2 \\ -pB_G^- \mathrm{diag}[\alpha_{ij}] B_G^T & \mathbf{0}_{n\times n} & -pN_2 \end{bmatrix}.$$

Thus, it is straight forward to see $(\hat{A} + \hat{B}_2 K \hat{C}_2) = NA$. Analogously we get $\hat{B}_1 + \hat{B}_2 K \hat{D}_{21} = NB$ and $\hat{C}_1 = C$. □

The main difficulty in solving (3.41) instead of (3.42) is that, as we showed

in Proposition 3.1, our controller K is a nonlinear function of the parameters $K(\kappa_1, \kappa_2, p, \alpha)$ and cannot be readily obtained using (3.42). Furthermore, Proposition 3.1 also shows that the main source of nonlinearity comes from the products $\kappa_1 \mathrm{diag}[\alpha_{ij}]$ and $p\mathrm{diag}[\alpha_{ij}]$. This structure is not currently supported by traditional software distributions, which tend to support only sparsity patterns, and therefore needs to be implemented.

One particular package that proved to be easily adapted was Hifoo [151,153] and more precisely in its discrete-time version Hifood [154]. These algorithms only use gradient information in their implementation of BGS and gradient bundle stages. Thus, to implement discrete time $\mathcal{H}_2$ optimization, a new Matlab subroutine that evaluated the $\mathcal{H}_2$ norm f as well as its gradient was created.

The evaluation of the gradient is performed in three stages using the chain rule. We first compute the gradients of f with respect to $\bar{A} := \hat{A} + \hat{B}_2 K \hat{C}_2$, $\bar{B} := \hat{B}_1 + \hat{B}_2 K \hat{D}_{21}$ and $\bar{C} := \hat{C}_1$ which are given by

$$\nabla_{\bar{A}} f = \frac{1}{f} X \bar{A} Y, \quad \nabla_{\bar{B}} f = \frac{1}{f} X \bar{B} \quad \text{and} \quad \nabla_{\bar{C}} f = \frac{1}{f} \bar{C} Y.$$

Once $\nabla_{\bar{A}} f$, $\nabla_{\bar{B}} f$ and $\nabla_{\bar{C}} f$ are computed we can use the subroutines of hifood to compute $\frac{\partial \bar{A}}{\partial K}$, $\frac{\partial \bar{B}}{\partial K}$ and $\frac{\partial \bar{C}}{\partial K}$. Finally, since $\frac{\partial K}{\partial \kappa_1}$ as well as the other parameters' derivatives can be computed using Proposition 3.1 we obtain

$$\nabla_{\kappa_1} f = \\ \mathrm{trace}\left[\left(\nabla_{\bar{A}} f^T \frac{\partial \bar{A}}{\partial K} + \nabla_{\bar{B}} f^T \frac{\partial \bar{B}}{\partial K} + \nabla_{\bar{C}} f^T \frac{\partial \bar{C}}{\partial K}\right) \frac{\partial K}{\partial \kappa_1}\right]$$

and similarly for other parameters.

3.5 Experiments

To test our solution and analysis, we implement an asynchronous version of our algorithm in C using the IBM CCT solution as our code base. Our program reads the TSC counter directly using the `rdtsc` assembly instruction to minimize reading latencies and maintains a virtual clock that can be directly updated. The list of neighbors is read from a configuration file and whenever there is no neighbor, the program follows the local Linux clock. Finally, offset measurements are taken using an improved ping pong mechanism proposed in [49].

We run our skewless protocol in a cluster of IBM BladeCenter LS21 servers with two AMD Opteron processors of 2.40GHz, and 16GB of memory. As shown in Figure 3.2, the servers serv1-serv10 are used to run the protocol. The offset measurements are taken through a Gigabit Ethernet switch. Server serv0 is used as a reference node and gathers time information from the different nodes using a Cisco 4x InfiniBand Switch that supports up to 10Gbps between any two ports and up to 240Gbps of aggregate bandwidth. This minimizes the error induced by the data collecting process.

We use this testbed to validate the analysis in Section 3.3. Firstly, we illustrate the effect of different parameters and analyze the effect of the network configuration on convergence (Experiment 1). Secondly, we present a series of configurations that demonstrate how connectivity between clients is useful in reducing the jitter of a noisy clock source (Experiment 2). Thirdly, we compare the performance of the algorithm with respect to NTP version 4 (Experiment 3) and a software-base version of IBM CCT (Experiment 4). Finally, we verify the

constant drift effect of path asymmetries predicted by Theorem 3.4 (Experiment 5) and verify the dependence of the optimal parameter values on the network topology and noise (Experiment 6).

We will use several performance metrics to evaluate our algorithm. The output performance signal v_k will be the vector of offset difference between the leader 1 and every other node i, i.e. $v_i(t_k) = x_i(t_k) - x_1(t_k)$ with $i \in \{2,...,n\}$, and use a normalized version of it herein mentioned as *mean relative deviation* , $\sqrt{S_n}$, as performance metric. In other words,

$$S_n = \frac{||v_k||_2^2}{n-1} = \frac{1}{n-1}\sum_{i=2}^{n}\left\langle (x_i - x_1)^2 \right\rangle. \tag{3.43}$$

where $< \cdot >$ amounts to the sample average. We will also use the 99% Confidence Interval CI_{99} and the maximum offset (CI_{100}) as metrics of accuracy. For example, if $CI_{99} = 10\mu s$, then the 99% of the offset samples will be within $10\mu s$ of the leader .

Unless explicitly stated, the default parameter values are

$$p = 0.99, \quad \kappa_1 = 1.1, \quad \kappa_2 = 1.0 \text{ and } \alpha_{ij} = \frac{c_i}{|\mathcal{N}_i|}. \tag{3.44}$$

The scalar c_i is a commit or gain factor that will allow us to compensate the effect of τ since $\alpha_{ii} = c_i$ for every node that is not the leader.

Notice that these values immediately satisfy (i) and (ii) of Theorem 3.3 since $1 - p = 0.01$, $\frac{2\kappa_1}{3p} = 0.7407 > \kappa_1 - \kappa_2 = 0.1$. The remaining condition can be satisfied by modifying τ or equivalently c. Here, we choose to fix $c_i = 0.7$ which makes condition (iii)

$$\tau < \frac{1.2717}{\mu_{\max}}\text{s}.$$

For fixed time step τ, the stability of the system depends on the value of $\mu_{\max}$, which is determined by the underlying network topology.

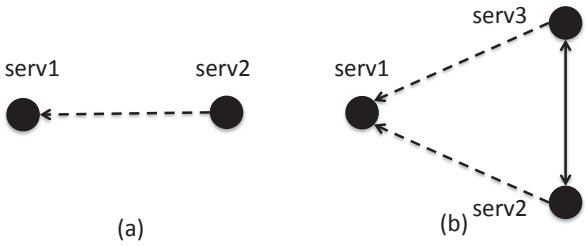
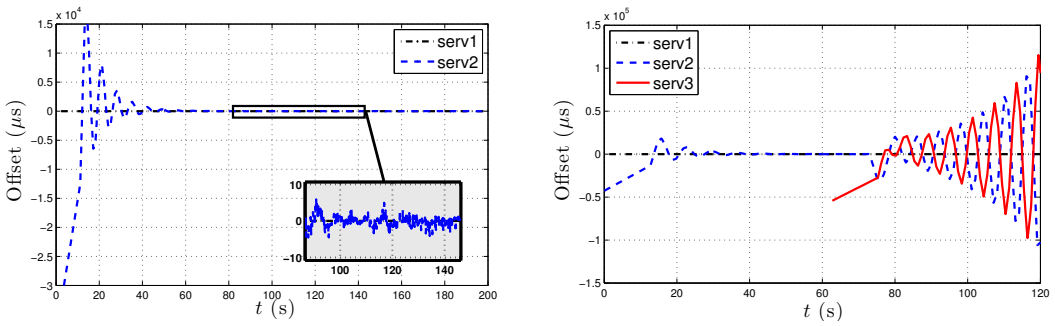


Figure 3.6: Effect of topology on convergence: (a) Client-server configuration; (b) Two clients connected to server and mutually connected.

Experiment 1 (Convergence): We first consider the client server configuration described in Figure 3.6a with a time step

$$\tau = 1\text{s}. \tag{3.45}$$

In this configuration $\mu_{\max} \approx 1$ and therefore condition (iii) becomes $\tau < 1.2717$s. Figure 3.7(a) shows the offset between serv1 (the leader) and serv2 (the client) in microseconds. There we can see how serv2 gradually updates s_2 until the offset becomes negligible for the plot scale.



(a) Client server configuration with $\tau = 1$s. The client converges and algorithm is stable. (b) Two clients mutually connected with $\tau = 1$s. The algorithm becomes unstable.

Figure 3.7: Lost of stability by change in the network topology

Figure 3.7(a) tends to suggest that the set of parameters given by (3.44) and (3.45) are suitable for deployment on the servers. This is in fact true provided that network is a directed tree as in Figure 3.5a. The intuition behind this fact is

that in a tree, each client connects only to one server. Thus, those connected to the leader will synchronize first and then subsequent layers will follow.

However, once loops appear in the network there is no longer a clear dependency since two given nodes can mutually get information from each other. This type of dependency might make the algorithm unstable. Figure 3.7(b) shows an experiment with the same configuration as Figure 3.7(a) in which serv2 synchronizes with serv1 until a third server (serv3) appears after 60s. At that moment the system is reconfigured to have the topology of Figure 3.6b introducing a timing loop between serv2 and serv3. This timing loop makes the system unstable.

The instability arises since after serv3 starts, the new topology has $\mu_{\max} \approx 1.5$. Thus, the time step condition (iii) becomes $\tau < 847.8$ms which is no longer satisfied by $\tau = 1$s.

Using (3.25) we can recover the stability of the system by setting

$$\tau = 500\text{ms} < \frac{1.2717}{2}\text{s} = 645.85\text{ms}$$

Figure 3.8 shows how serv2 and serv3 can now synchronize with serv1 after introducing this change.

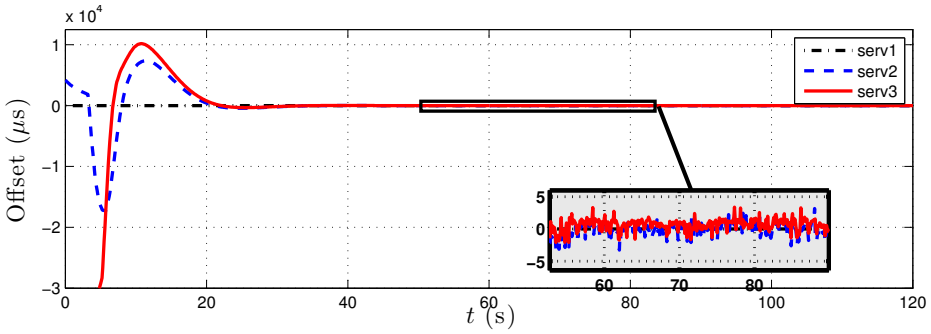


Figure 3.8: Two clients mutually connected with $\tau = 500$ms

Experiment 2 (Timing Loops Effect): We now show how timing loops can be used to collectively outperform individual clients when the time source is noisy (jitter).

We run our algorithm on 10 servers (serv1 through serv10). The connection setup is described in Figure 3.9. Every node is directly connected unidirectionally to the leader (serv1) and bidirectionally to $2K$ additional neighbors. When

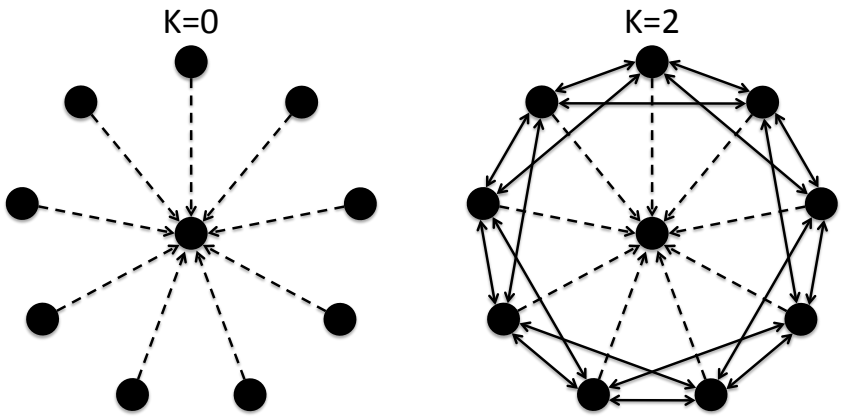


Figure 3.9: Leader topologies with $2K$ neighbors connection. Connections to the leader (serv1) are unidirectional while the connections among clients (serv2 trhough serv10) are bidirectional

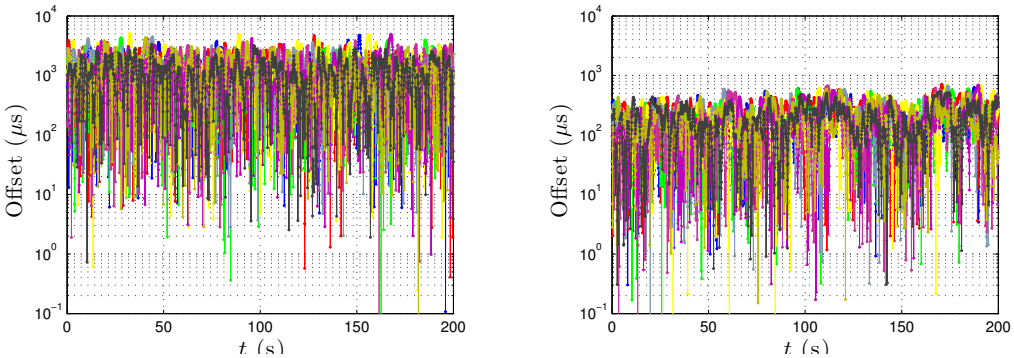
$K = 0$ the network reduces to a star topology and when $K = 4$ the servers serv2 through serv10 form a complete graph.

The dashed arrows in Figure 3.9 show the connections where jitter was introduced. To emulate a link with jitter, we added random noise η with values taken uniformly from $\{0, 1, ..., \text{Jitter}_{\max}\}$ on both directions of the communication,

$$\eta \in \{0, 1, ..., \text{Jitter}_{\max}\}\text{ms}. \tag{3.46}$$

Notice that the arrow only shows a dependency relationship, the ping pong mechanism sends packets in both directions of the physical communication [49].

We used a value of $\text{Jitter}_{\max} = 10$ms. Since the error was introduced in both directions of the ping pong, this is equivalent to a standard deviation of 4.69ms[3].



(a) Star topology ($K = 0$)

(b) Complete subgraph ($K = 4$)

Figure 3.10: Offset of the nine servers connected to a noisy clock source

Figure 3.10 illustrates the relative offset between the two extreme cases; The star topology ($K = 0$) is shown in Figure 3.10(a), and the complete subgraph ($K = 4$) is shown in Figure 3.10(b).

The worst case offset for $K = 0$ is 5.1ms which is on the order of the standard deviation of the jitter. However, when $K = 4$ we obtain a worst case offset of 690.8μs, an order of magnitude improvement.

The mean relative deviation $\sqrt{S_n}$ as the connectivity among clients increases from isolated nodes ($K = 0$) to a complete subgraph ($K = 4$) is studied in Figure 3.11. The results presented show that without any type of error filtering the network itself is able to perform a distributed filtering that achieves an improvement of up to a factor of 6.26 or equivalently a noise reduction of almost 8dB.

Experiment 3 (Comparison with NTPv4): We now perform a thorough com-

[3]The value 4.69ms is the standard deviation of the sum of two uniform distributed random variables.

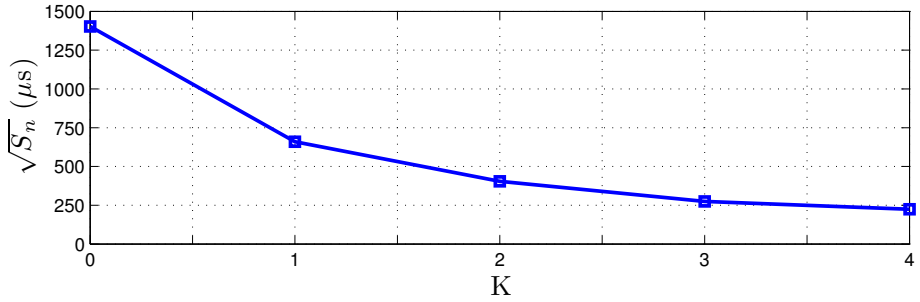


Figure 3.11: Effect of the client's communication topology on the mean relative deviation. As the connectivity increases (K increases) the mean relative deviation is reduced by factor of 6.26, i.e. a noise reduction of approx. 8dB.

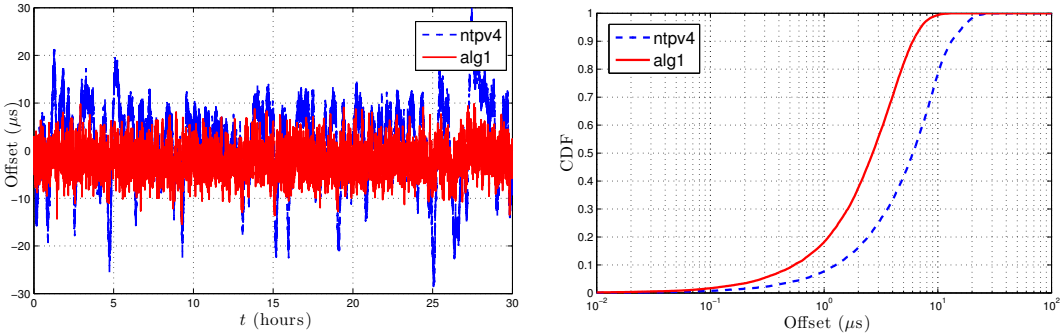
parison between our algorithm (Alg1) and NTPv4. We will use the one hop configuration of Figure 3.6b but without the bidirectional link between serv2 and serv3. Here, server serv1 is set as NTP server and as leader of Alg1, server serv2 has a client running NTPv4 and server serv3 a client running our algorithm.

In order to make a fair comparison, we need both algorithms to use the same polling interval. Thus, we fix $\tau = 16$sec. This can be done for NTP by setting the parameters `minpoll` and `maxpoll` to 4 ($2^4 = 16$secs). The remainder parameter values for our algorithm are given by

$$p = 1.98, \quad \kappa_1 = 1.388 \text{ and } \kappa_2 = 1.374. \tag{3.47}$$

Figure 3.12(a) shows the time differences between the clients running NTPv4 and Alg1 (serv2 and serv3), and the leader (serv1) over a period of 30 hours. It can be seen that Alg1 is able to track serv1's clock keeping a offset smaller than 10μs for most of the time while NTPv4 incurs in larger offsets during the same period of time. This difference is produced by the fact that Alg1 is able to react more rapidly to frequency changes while NTPv4 incurs in more offset

corrections that generate larger jitter.



(a) Offset values of NTPv4 and Alg1 for a period of 30 hours.

(b) Cummulative Distribution Function

Figure 3.12: Performance evaluation between our solution (Alg1) and NTPv4

A more detailed and comprehensive analysis is presented in Figure 3.12(b) where we plot the Cumulative Distribution Function (CDF) of the offset samples. That is, the fraction of samples whose time offset is smaller than a specific value. Using Figure 3.12(b) we compute the corresponding 99% confidence intervals (CI_{99})

Our algorithm (Alg1) achieves a performance with $\sqrt{S_n} = 3.1\mu$s, $CI_{99} = 9.5\mu$s and a maximum offset of 15.9μs, while NTPv4 obtains $\sqrt{S_n} = 8.1\mu$s, $CI_{99} = 21.8\mu$s and a maximum offset of 28.0μs. Thus, not only Alg1 achieves a reduction of $\sqrt{S_n}$ by a factor of 2.6 (-4.2dB) with respect to NTPv4, but it also obtains smaller confidence intervals and maximum offset values.

Finally, we investigate the speed of convergence. Starting from both clients synchronized to server serv1, we introduce a 25ms offset. Figure 3.13 shows how Alg1 is able to converge to a $20\mu s$ range within one hour while NTPv4 needs 4.5hours to achieve the same synchronization precision. In summary, not only can our algorithm achieve better performance than NTPv4, but it can also

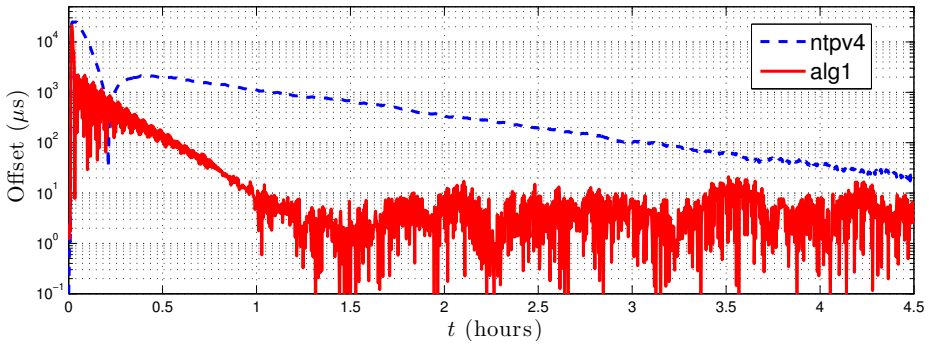
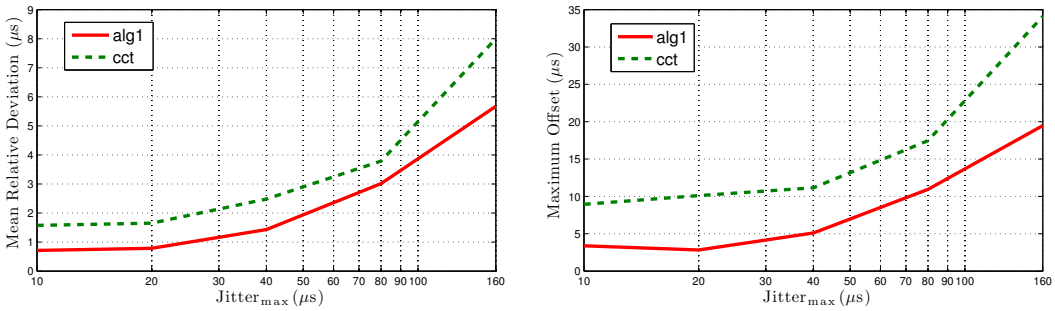


Figure 3.13: Offset values of NTPv4 and Alg1 after a 25ms offset introduced in serv1.

converge faster.

Experiment 4 (Comparison with IBM CCT): We now proceed to compare the performance of our algorithm (Alg1) with respect to IBM CCT. Notice that unlike IBM CCT, our algorithm does not perform any previous filtering of the offset sample, the filtering is performed instead by calibrating the parameters which mostly depend on the polling interval τ chosen. Here we use $c_i = 0.70$, $\tau = 250ms$, $\kappa_1 = 0.1385$, $\kappa_2 = 0.1363$ and $p = 0.62$.



(a) Mean relative deviation $\sqrt{S_n}$ (b) Maximum offset

Figure 3.14: Performance evaluation between our solution (Alg1) and IBM CCT

In Figure 3.14(a) we present the mean relative deviation $\sqrt{S_n}$ for two clients connected directly to the leader as the jitter is increased from $\text{Jitter}_{\max} = 0\mu$s

(no jitter) to $\text{Jitter}_{\max} = 160\mu$s with a granularity of 1μs. The worst case offset is shown in Figure 3.14(b). Each data point is computed using a sample run of 250 seconds.

Our algorithm consistently outperforms IBM CCT in terms of both $\sqrt{S_n}$ and worst case offset. The performance improvement is due to two reasons. Firstly, the noise filter used by the IBM CCT algorithm is tailored for noise distributions that are mostly concentrated close to zero with sporadic large errors. However, it does not work properly in cases where the distribution is more homogeneous as in this case. Secondly, by choosing $\delta\kappa = \kappa_1 - \kappa_2 = 0.002 \ll 1$ and the discussion in Section (3.2) we can see that κ in (3.14) becomes very small, which makes the algorithm more sensitive to frequency mismatches than offsets. This makes the algorithm very robust to offset errors.

Experiment 5 (Frequency drift without leader): We now proceed to experimentally verify that without leader, the system tends to constantly drift the frequency. Our analysis predicts that even the minor bias in the offset measurements will produce this effect. To verify this phenomenon, we use the network topology in Figure 3.6b with $\tau = 0.5$s and wait for the system to converge.

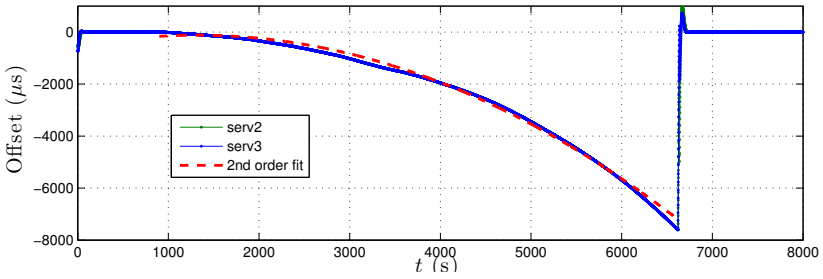


Figure 3.15: Frequency drift

After 1000s the timing process of serv1 is turned off. Figure 3.15 shows how

the offsets of serv2 and serv3 start to grow in a parabolic trajectory characteristic of a constant acceleration, i.e. constant drift. After 6600s serv1 is restarted and the system quickly recovers synchronization. A second order fit of the faulty trajectory was performed obtaining a drift of approximately -250 ns/s^2. While this is not quite significant in the first few minutes, it becomes significant as time goes on.

Experiment 6 (Jitter and Wander Tradeoff): Finally, we use the proposed $\mathcal{H}_2$ optimization scheme to show how the optimal parameter values depend on the different noise condition within the network described in Figure 3.16. We consider three different noise scenarios in which we either add jitter between server serv1 and servers serv2 and serv3, and/or add wander on severs serv2-serv7. In all the cases we use $\tau = 0.5$s and make offset measurements through the InfiniBand switch to minimize the any additional source of noise.

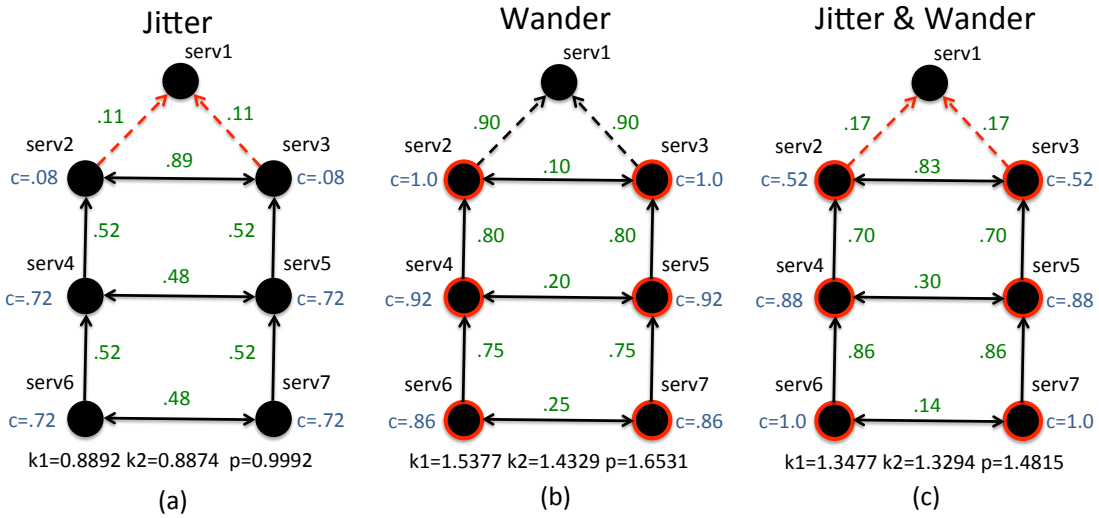


Figure 3.16: Network scenarios and optimal parameters

The jitter is generated by adding in both directions of the physical com-

munication a random value η similarly to Experiment 2(c.f. (3.46)), but with a $\text{Jitter}_{\max} = 100\mu$s. This generates an aggregate offset measurement noise of zero mean and standard deviation of 40.8μs. On the other hand, the wander is generated by adding gaussian noise with zero mean and standard deviation of 0.2ppm in the $s_i(t_k)$ adaptations. As discussed in Section 3.4, this noise can be used to emulate the wander of a bad quality clock.

We used different values of g_{ij}^w and g_i^d to differentiate the noise conditions in the optimization scheme. The large jitter scenario is represented by $g_i^d = 1e{-}3\ \forall i$, $g_{21}^w = g_{31}^w = 100$ and $g_{ij}^w = 1$ otherwise. The large wander scenario is represented by $g_i^d = 1e-1\ \forall i$ and $g_{ij}^w = 1$. Finally, the large jitter and wander scenario is represented using $g_i^d = 1e-1\ \forall i$, $g_{21}^w = g_{31}^w = 100$ and $g_{ij}^w = 1$ otherwise. The output parameter values for all three cases are also present in Figure 3.16.

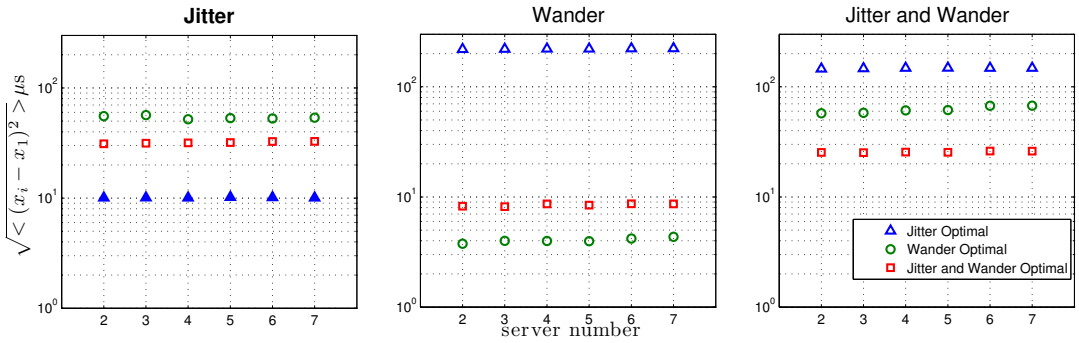


Figure 3.17: $\mathcal{H}_2$ Performance optimization: offset variance vs server number

Figure 3.17 shows the standard deviation of the offset between servers serv2-serv7 and serv1 in the three experimental scenarios and for the three different sets of parameters shown in Figure 3.16. It can be seen that although the configuration tuned for jitter performs very well in cases with large jitter, it performs quite poorly in scenarios with large wander. Similarly, the configuration tuned

for wander does not perform well in high jitter scenarios.

However, the configuration tuned for jitter and wander is able to provide acceptable performance in all three experimental scenarios. Thus, we experimentally demonstrate a fundamental tradeoff between offset and wander.

CHAPTER 4

SYNCHRONIZATION ON POWER NETWORKS

In this chapter we focus on the study of the synchronization of a power grid and how its performance is affected by the different conditions of the network. Using a local stability analysis similar to the one conducted for coupled oscillators in section 2.4, we relate the damping of the network with the different network parameters and provide an updating direction that decreases it. The analysis suggests that one can use power scheduling or modify line impedances in order to prevent saddle-node bifurcations. However, this result triggers more questions than answers. Firstly, many of these parameters are usually set using the output of an OPF with a given economic performance objective. Secondly, even if it is possible to include the damping as part of the OPF problem, it is not even clear that this metric is suitable to measure the stability of a power grid.

In this chapter we shall answer all these questions. In section 4.1 we describe the dynamics of a power network, the different stability issues it can experience and the standard OPF problem. We also describe a simplified model in section 4.1.3, closely related to coupled oscillators, that will be key in understanding the interplay between network parameters and stability. Section 4.2 then relates the damping of a power network with the second smallest eigenvalue of a state dependent weighted Laplacian. We then characterize the dependence of the eigenvalue, a.k.a. algebraic connectivity, of this Laplacian in term of its weights in section 4.3, and derive updating directions that improve the damping of a network in 4.3.1 and 4.3.2. We illustrate our findings using numerical examples in section 4.5.1.

We then focus on understanding which performance metric is more efficient

in characterizing the system's stability. With this aim, we bring in the pseudospectral abscissa in section 4.4 and show how it can be used to measure and optimize voltage stability margins, oscillations and robustness. This naturally leads to our Dynamics-aware OPF formulation. Finally, we illustrate several properties of our new optimization framework using two different test cases, including the widely used IEEE 39-bus New England power grid test case in section 4.5.2.

4.1 Power Network Modeling

We now proceed to describe two models commonly used in the study of OPF and power system dynamics: *static* and *dynamic* models. Each one has its specific use and the level of detail depends on the problem in consideration.

4.1.1 Static Model

The *static* model of a power network defines the physical relationship that the state at each bus must satisfy for the system to be at equilibrium. In this model, the state is solely represented by the complex voltage $V_i = |V_i|e^{j\theta_i}$ at each bus $i \in V$, which in order to be at equilibrium, must satisfy the flow conservation equations, also known as *power flow equations*. These equations basically state that the surplus (or deficit) in generation at a given bus should match the outgoing (incoming) power flow to (from) the neighboring buses and ground, i.e.

$$|V_i|^2 y_{ii}^* + \sum_{j \in \mathcal{N}_i} S_{ij} = P_{G_i} + jQ_{G_i} - (P_{D_i} + jQ_{D_i}). \tag{4.1}$$

Here, $P_{G_i} + jQ_{G_i}$ is the complex power generated, $P_{D_i} + jQ_{D_i}$ is the complex power demanded at bus i, $S_{ij} = P_{ij} + jQ_{ij} := V_i(V_i - V_j)^* y_{ij}^*$ is the complex line flow from i to j, y_{ii} is the bus shunt admittance and $y_{ij} := g_{ij} + jb_{ij}$ is the line admittance. Loads are usually modeled as constant impedance (Z), constant current (I) or constant power (P). When the loads are modeled by constant impedance or constant current models, P_{D_i} and Q_{D_i} are functions of the voltage magnitude at the bus. A well-accepted model for static loads is the ZIP model which is a convex combination of the three, i.e.

$$P_{D_i} = P_{0,i}\left(a_{1,i}\left(\frac{|V_i|}{V_{0,i}}\right)^2 + a_{2,i}\left(\frac{|V_i|}{V_{0,i}}\right) + a_{3,i}\right) \tag{4.2a}$$

$$Q_{D_i} = Q_{0,i}\left(b_{1,i}\left(\frac{|V_i|}{V_{0,i}}\right)^2 + b_{2,i}\left(\frac{|V_i|}{V_{0,i}}\right) + b_{3,i}\right) \tag{4.2b}$$

Since this model is sufficient to characterize the static properties of the network, such as the existence of a stationary solution of the power flow equations (4.1), voltage magnitudes $|V_i|$, line flows P_{ij} and S_{ij}, and losses $P_{ij} + P_{ji}$, it is used for the computation of the optimal power flow and the study of static voltage stability.

To simplify notation, we will use from now on $x_s := [|V|^T\ \theta^T]^T$ as the vector of the static network states, $u_s := [P_G^T\ Q_G^T]^T$ as the vector of static control variables and $v_s := [P_0^T\ a_1^T\ a_2^T\ a_3^T\ Q_0^T\ b_1^T\ b_2^T\ b_3^T]^T$ as the vector of load parameters. Thus, the power flow equations (4.1) can be compactly defined as $F(x_s, u_s, v_s) = 0$.

Optimal Power Flow

Let $f_i(V_i, P_{G_i}, Q_{G_i})$ denote the cost function associated with bus i. In most cases, f_i depends solely on P_{G_i} but it can be extended to more general scenarios. Then,

the optimal power flow can be formulated as

$$\textbf{OPF} : \underset{x_s, u_s}{\text{minimize}} \quad c(V, P_G, Q_G) := \sum_{k \in \mathcal{N}} f_i(V_i, P_{G_i}, Q_{G_i}) \tag{4.3}$$

subject to

$$F(x_s, u_s, v_s) = 0 \tag{4.4a}$$

$$P_i^{\min} \leq P_{G_i} \leq P_i^{\max}, \qquad \forall i \in \mathcal{N} \tag{4.4b}$$

$$Q_i^{\min} \leq Q_{G_i} \leq Q_i^{\max}, \qquad \forall i \in \mathcal{N} \tag{4.4c}$$

$$V_i^{\min} \leq |V_{ij}| \leq V_i^{\max}, \qquad \forall i \in \mathcal{N} \tag{4.4d}$$

$$P_{ij} \leq P_{ij}^{\max}, \qquad \forall ij \in \mathcal{L} \tag{4.4e}$$

$$|S_{ij}| \leq S_{ij}^{\max}, \qquad \forall ij \in \mathcal{L} \tag{4.4f}$$

The list of methods to solve this problem is vast. Some of the most commonly used are primal dual interior point method [94], trust region based augmented Lagrangian [95], newton method [96] and successive linear programming [97].

Voltage Stability

Voltage stability refers to the ability of the system to preserve voltage magnitudes within its nominal values and avoid voltage collapse. A voltage collapse occurs when changes on u_s or v_s make two solutions of (4.4a) coalesce and disappear in a **Saddle Node Bifurcation**. This is evidenced by the presence of a real eigenvalue of the Jacobian matrix

$$J(x_s, u_s, v_s) = D_{x_s} F(x_s, u_s, v_s) \tag{4.5}$$

on the imaginary axis.

It is important to notice that the OPF problem (4.3)-(4.4) guarantees voltage stability since its solution satisfies the power flow constraints (4.4a). However, the stability margins may not be large and a small fluctuation on the demand can thus produce a voltage collapse.

This has motivated the development of optimization-based techniques that define some distance measure, compute the smallest distance to voltage collapse (e.g. [67, 155]) and improve it [66, 68, 156, 157]. These developments have led to a solid integration of voltage stability measures as constraints or as part of the objective function of the OPF problem [103–109]. Yet, none of them considers the effect of the outcome of these solutions on the dynamics of the power system.

4.1.2 Dynamic Model

The dynamics of a power network are represented by a set of differential algebraic equations (DAEs) [158]

$$\dot{x} = f(x, z, u, v) \tag{4.6a}$$

$$0 = g(x, z, u, v). \tag{4.6b}$$

where x and z are the slow and fast state variables, u are the control inputs, such as power generation, active voltage regulators (AVR) set points, transformers taps, etc., and v are the exogenous parameters such as power demand. Equation (4.6a) represents the dynamics of the system devices, including generators, power electronics and controllers, and (4.6b) are the algebraic equations of the generators stators, power electronics and network power flows.

Equations (4.6a)-(4.6b) form a more detailed model than the static

model(4.1)-(4.2) and include in (x, z), u and v, the values of x_s, u_s and v_s, respectively. In fact, equation (4.4a) is a subset of (4.6a)-(4.6b).

Remark 4.1. *It is important to notice that when x_s, u_s and v_s satisfy $F(x_s, u_s, v_s) = 0$, we can find x, z such that $f(x, z, u, v) = 0$ and $g(x, z, u, v) = 0$. This will be used in later sections to formulate our Dynamics-aware OPF. Overall, the level of detail in the dynamic model is essential when one wants to study dynamic phenomena such as small signal oscillations.*

Small Signal Oscillations

Small signal oscillations are the effect of a **Hopf Bifurcation** in which a stable equilibrium point becomes unstable and a limit cycle appears, or the effect of poorly damped modes of stable operating points. These oscillations can be studied by linearizing the system (4.6a)-(4.6b) around an equilibrium point (x^*, z^*, u^*, v)

$$\dot{x} = [D_x f]x + [D_z f]z + [D_u f]u \tag{4.7a}$$

$$0 = [D_x g]x + [D_z g]z + [D_u g]u \tag{4.7b}$$

and assuming that $D_z g(x^*, z^*, u^*, v)$ is nonsingular[1] to obtain reduced system

$$\dot{x} = Ax + Bu \tag{4.8}$$

where

$$A = \left[D_x f - D_z f \left(D_z g\right)^{-1} D_x g\right](x^*, z^*, u^*, v) \tag{4.9}$$

and

$$B = \left[D_u f - D_z f \left(D_z g\right)^{-1} D_u g\right](x^*, z^*, u^*, v).$$

[1]The nonsingularity of $D_z g(x^*, z^*, u^*, v)$ is a standard assumption in power system stability studies that is generally satisfied, see e.g. [159].

The presence of small signal oscillations is evidenced by the presence of a complex conjugate pair of eigenvalues of A close to the imaginary axis. As previously mentioned, small signal stability can usually be improved by designing controllers (e.g. PSS and FACTS) such that in closed loop A has eigenvalues with smaller damping ratio [70, 71, 80–86, 86–89]. However, none of these solutions considers the fact that (4.9) depends on the solution of the power scheduling (encoded in u^*) and that oscillations can appear if the market solution moves the system towards a more stressed condition. This generates the need for redispatching procedures that correct the scheduling in order to avoid small signal instabilities.

The current way of dealing with the above issue is by either iteratively adding constraints to successive OPF instances based on eigenvalues sensitivity information [102,110,111] or solving an OPF instance using an interior point method with a constraint on the real part $\Re[\lambda_i]$ of every critical eigenvalue [112]. Besides the computational complexity of these methods (one of them has to solve several OPFs and the others compute second order sensitivity of eigenvalues), it is also important to notice that most of them essentially use $\max \Re[\lambda_i]$ as a stability constraint to avoid Hopf Bifurcations, and disregard any other performance or robustness metric in the optimization. The only exception is [102] which successively adds approximate damping ratio constraints to each OPF instance solved. Using the function $\max \Re[\lambda_i]$ as stability measure is undesirable because it can make the system exhibit late amplitude oscillations as one gets closer to a local minimum of it [122,123]. On the other hand, adding damping ratio constraints on the eigenvalues has no effect on voltage stability, as a real eigenvalue can be arbitrarily close to the imaginary axis without meeting any damping constraint. These difficulties directly motivates us to formulate a

Dynamics-aware OPF.

4.1.3 Network Preserving Dynamic Model

Finally, we describe a simplified version of (4.6) that was first introduced by Bergen and Hill in 1981 [119]. The Bergen-Hill model is derived by making several simplifying assumptions:

1. *Lossless:* Every transmission line has zero conductance, i.e. $y_{ij} = jb_{ij}$.
2. *Decoupling:* The power flow equations (4.1) can be decoupled such that the phases ϕ_i depend only on P_i and the voltage magnitudes $|V_i|$ depend on Q_i.
3. *Load model:* Loads are modeled assuming constant reactive power $Q_{D_i} := Q^0_{D_i}$ and frequency dependent real power $P_{D_i}(\dot{\phi}_i) := P^0_{D_i} + D_i\dot{\phi}_i$.
4. *Generator model:* Generators are modeled by a constant internal voltage and transient reactance with swing dynamics.

While assumption 1 is very common in the literature and is not necessarily critical, assumption 2 and 3 together have a significant impact on the model. For example, since by assumption 3 the reactive power is constant, assumption 2 implies that the voltage magnitude at every bus is constant too. This allows us to eliminate the imaginary part of equation (4.1), which together with assumption 1 gives

$$P_i := P_{G_i} - P_{D_i} = \sum_{l \in \mathcal{N}_i} |V_i||V_j| b_{ij} \sin(\phi_i - \phi_j),$$

with P_i being the power injection at bus k.

Finally, assumption 4 allows us to substitute the generator with a constant voltage internal bus with a lossless transmission line. Thus, we can completely describe the state of each generator using ϕ_i and $\omega_i = \dot{\phi}_i$ which evolve according to

$$M_i\ddot{\phi}_i + D_i\dot{\phi}_i = P_{G_i} - P_{e_i} \quad \forall k \in \{1,...,m\}.$$

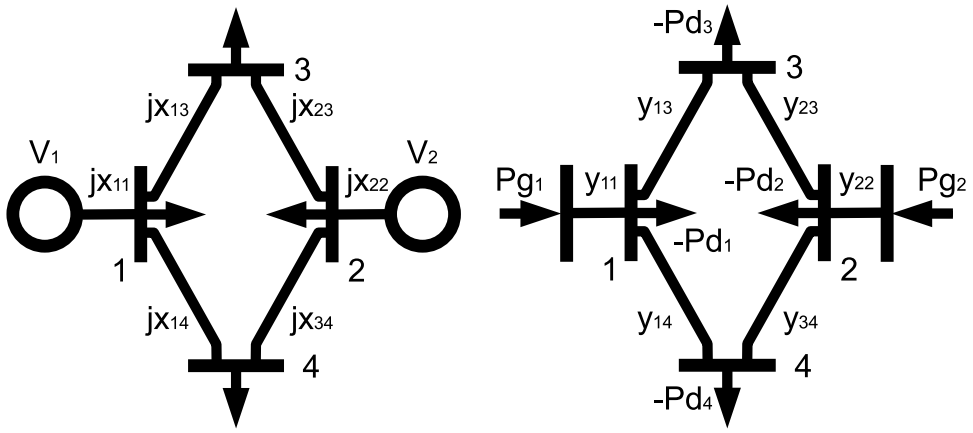
where M_i and D_i are the generator's inertia and damping, P_{G_i} is the mechanical power, P_{e_i} is the electrical real power that the network is demanding from the generator and m is the number of generators.

Thus, given a network composed by n buses, we obtain an extended network with m generator buses plus n load buses whose dynamics are described by

$$M_i\ddot{\phi}_i + D_i\dot{\phi}_i = P_i + \sum_{l\in\mathcal{N}_i} |V_i||V_j|b_{ij}\sin(\phi_i - \phi_j) \quad i \in \{1,...,m\} \tag{4.10a}$$

$$D_i\dot{\phi}_i = P_i + \sum_{l\in\mathcal{N}_i} |V_i||V_j|b_{ij}\sin(\phi_i - \phi_j) \quad i \in \{m+1,...,m+n\} \tag{4.10b}$$

See Figure 4.1(a) for an illustration of a sample power network with four buses and two generators.



(a) Real Power Network (b) Equivalent Network

Figure 4.1: Power Network Representations

Equation (4.10) can also be compactly expressed in vector form using

$$M\ddot{\phi} + D\dot{\phi} = -B\left(f(b) \circ \sin(B^T\phi)\right) + P, \tag{4.11}$$

where $(a \circ b)_i = a_i b_i$ is the Hadamard product between the vectors a and b, and $(f(b))_{ij} = |V_i||V_j|b_{ij}$ is the maximum instantaneous power flow between i and j. The diagonal matrices M and D represent the generators' inertia, generators' damping and loads' frequency coefficients, i.e.

$$(M)_{ij} = \begin{cases} M_i, \text{ if } i = j \; i \in \{1,...,m\}, \\ 0 \text{ o.w.}, \end{cases} \qquad (D)_{ij} = \begin{cases} D_i \text{ if } i = j, \\ 0 \text{ o.w.}. \end{cases}$$

We will use D_{max}, D_{min} and M_{max} to denote the nonzero extreme values that D and M can reach. The vector $P \in \mathbb{R}^{m+n}$ is the power injection at each bus, i.e.

$$P_i = \begin{cases} P_{G_i} & \forall i \in \{1,...,m\} \\ -P_{D_i} & \forall i \in \{m+1,...,m+n\}. \end{cases}$$

And the matrix $B \in \mathbb{R}^{(m+n)\times(m+n)(m+n-1)}$ is the incidence matrix of the complete graph.

Remark 4.2. *The matrix B as defined here does not capture (alone) the topology of the network. It is the conjunction of B and $b = (b_{ij})$ that captures the topology since $b_{ij} > 0$ if and only if ij represent a line of the extended power network graph $G = (V, E)$, i.e. iff $ij \in E$. In this way the addition of a line does not change the dimension of B.*

4.2 Effect of Topology

The damping of (4.11) can be locally estimated by computing the eigenvalues of the Jacobian J_{ϕ^*} of the linearized version of (4.11) around a given equilibrium

$(\phi^*, \dot{\phi}^* = 0)$,

$$M\delta\ddot{\phi} + D\delta\dot{\phi} + L(w(\phi^*))\delta\phi = 0, \tag{4.12}$$

where the matrix $L(w(\phi^*)) := B\mathrm{diag}[w(\phi^*)]B^T$ represents the weighted Laplacian of the graph G with weights $w_{ij}(\phi^*) = |V_i|\left|V_j\right| b_{ij}\cos(\phi_j^* - \phi_i^*)$, and captures several topological properties of the network (see e.g. [120]).

When $(\phi^*, 0)$ is stable, $L(w(\phi^*))$ is positive semidefinite with, under generic conditions, only one zero eigenvalue $\nu_1(L(w(\phi^*)))$ with eigenvector $\mathbf{1}_{m+n}$. This implies that the smallest eigenvalue of (4.12), λ_1, is zero. The existence of this zero eigenvalue is due to the rotational symmetry of the system, $\sin(B^T(\phi + \mathbf{1}_{m+n})) = \sin(B^T\phi)$.

If $D_i \gg M_i$, it is possible to approximate (4.11) by setting $M = 0$ [160]. Then, (4.12) becomes a first order system and the damping can be upper-bounded by,

$$\Re[\lambda_2] \leq -\cos(\alpha^*)\frac{\nu_2(L(w(\phi^*)))}{D_{max}}, \tag{4.13}$$

where $\alpha^* := \angle(D\mathbf{1}, \mathbf{1})$ is the angle between vectors $D\mathbf{1}$ and $\mathbf{1}$; we use $\Re[\cdot]$ and $\Im[\cdot]$ to denote the real and imaginary part of a complex element.

Equation (4.13) suggests a correlation between $\Re[\lambda_2]$ and the algebraic connectivity $\nu_2(L(w(\phi^*)))$; however, a priori this relation seems to be only valid when $D_i \gg M_i$. The problem is that when $D_i \not\gg M_i$, the computation of the eigenvalues of (4.12) is usually done by introducing the state variables $\delta\omega = \delta\dot{\phi}$ and interpreting (4.12) as a first order linear system of dimension $2m + n$. This approach hides the rich symmetry inherent to M, D and $L(w(\phi^*))$ and makes the generalization of (4.13) hard.

In this thesis we use a more elegant approach to relate the damping of (4.11) with $\nu_2(L(w(\phi^*)))$. Using Matrix Polynomial Theory [161], we show that the

when the network is close to a bifurcation [162] (4.13) still holds. This approach is summarized next.

Instead of solving the linear eigenvalue problem of finding pairs $(\lambda_i, v_i) \in \mathbb{C} \times \mathbb{C}^{2m+n}$ such that $(\lambda_i I - J_{\phi^*})v_i = 0$, we solve the quadratic eigenvalue problem [163] of finding $(\lambda_i, x_i) \in \mathbb{C} \times \mathbb{C}^{m+n}$ such that $Q(\lambda_i)x_i = 0$ with $Q(\lambda) = M\lambda^2 + D\lambda + L(w(\phi^*))$. It is easy to show that $Q(\lambda)$ has $2(m+n)$ eigenvalues and that if λ_i is an eigenvalue of J_{ϕ^*}, it is also an eigenvalue of $Q(\lambda)$ [161]. The difference in the number of eigenvalues is due to the fact that M has n zero eigenvalues, which is reflected in $Q(\lambda)$ with the presence of n infinite eigenvalues. However, since we are only concerned about the dominant eigenvalue of (4.12), these infinite eigenvalues are not of interest to us.

This is a more natural formulation, since now the symmetry of M, D and $L(w(\phi^*))$ implies that if x_i is a right eigenvector of $Q(\lambda)$ then its complex conjugate $\bar{x}_i$ is a left eigenvector, and given the finite pair (λ_i, x_i) the following relationship holds

$$\lambda_i = \begin{cases} -\frac{l(x_i)}{d(x_i)} & \text{if } m(x_i) = 0, \\ \frac{-d(x_i) \pm \sqrt{d(x_i)^2 - 4m(x_i)l(x_i)}}{2m(x_i)} & \text{otherwise.} \end{cases} \tag{4.14}$$

where $m(x) = \bar{x}^T M x$, $d(x) = \bar{x}^T D x$ and $l(x) = \bar{x}^T L(w(\phi^*))x$. Notice that since $M \geq 0$, $D > 0$ and $L(w(\phi^*)) \geq 0$, $m(x)$, $d(x)$ and $l(x)$ are real and for any $x \neq 0$, $m(x) \geq 0$, $d(x) > 0$ and $l(x) \geq 0$.

The next theorem extends (4.13) to some cases where $M \neq 0$.

Theorem 4.1 (Damping Bound)**.** *When* (4.11) *is close to a bifurcation, the dominant eigenvalue* λ_2 *of* (4.12) *is real and bounded by* (4.13)*.*

Proof. Since the system is assumed to be close to a bifurcation, then $\Re[\lambda_2]$ must

close to the imaginary axis, i.e. $|Re[\lambda_2]| \ll \frac{D_{\min}}{M_{\max}}$. Thus, it follows from (4.14) that λ_2 is real ($\Im[\lambda_2] = 0$).

We first show that $\Re[\lambda_2] \leq -\frac{l(x_2)}{d(x_2)}$, which is trivial from (4.14) if $m(x_2) = 0$. Thus consider the case of $m(x_2) > 0$. Since $\Im[\lambda_2] = 0$, (4.14) implies $d(x_2)^2 > 4m(x_2)l(x_2)$. Thus, using the fact that $\sqrt{1-x} \leq 1 - \frac{1}{2}x$ and (4.14), we get

$$\Re[\lambda_2] \leq \frac{-d(x_2) + d(x_2)\left(1 - \frac{1}{2}(\frac{4m(x_2)l(x_2)}{d(x_2)^2})\right)}{2m(x_2)} = -\frac{l(x_2)}{d(x_2)}.$$

Therefore, whenever $\Im[\lambda_2] = 0$, $\Re[\lambda_2] = \lambda_2 \leq -\frac{l(x_2)}{d(x_2)}$.

The main problem with this bound is that since $l(x)$ is not positive definite, it cannot be readily lower bounded by a positive value. We therefore need to use the fact that x_2 is an eigenvector of $Q(\lambda)$ to obtain an appropriate lower bound on $l(x_2)$.

Since x_2 is an eigenvector and $\lambda_2 \neq 0$, then it follows from $\mathbf{1}_{m+n}^T L_{\phi^*} = 0$ and $Q(\lambda_2)x_2 = 0$ that,

$$\begin{aligned} 0 &= \mathbf{1}_{m+n}^T Q(\lambda_2)x_2 = \mathbf{1}_{m+n}^T(\lambda_2^2 M + \lambda_2 D + L_{\phi^*})x_2 \\ &= \mathbf{1}_{m+n}^T(\lambda M + D)x_2. \end{aligned}$$

So, when $\lambda_2 \ll \frac{D_{\min}}{M_{\max}}$, $x_2 \in \ker[\mathbf{1}_{m+n}^T D]$ and it follows that the biggest angle that x_2 can achieve with respect to $\ker(\mathbf{1}^T)$ is α^* and therefore $l(x_2) \geq \cos(\alpha^*)\|x_2\|^2 > 0$. Finally, since $d(x_2) \leq D_{max}\|x_2\|^2$, we get (4.13).

□

4.3 Improving Damping of a Stable Equilibrium

This section introduces some updating rules on the network parameters that improve the damping of a power network close to a saddle node bifurcation. In order to do this we need to know how the second smallest eigenvalue $\nu_2(L(w)) =: \nu_2(w)$ of the Laplacian $L(w)$ changes with w. There are several nice properties of $\nu_2(w)$ when $L(w)$ is positive semidefinite, see e.g. [120]. In particular, $\nu_2(w)$ is a concave function of w and homogeneous of degree one , i.e. $\nu_2(\lambda w) = \lambda \nu_2(w)$.

Here we are interested in computing $\frac{\partial}{\partial w_{ij}} \nu_2(w)$ whenever it is possible. If for given w the multiplicity of $\nu_2(w)$ is one, $\nabla_w \nu_2(w)$ is defined and can be readily computed by expressing $\nu_2(w)$ as

$$\begin{aligned} \nu_2(w) &= \min_{\{x:\|x\|=1,\langle x,\mathbf{1}\rangle=0\}} x^T L(w) x \\ &= \min_x \max_{\mu_1,\mu_2} W(w, x, \mu_1, \mu_2) \\ &= x^*(w)^T L(w) x^*(w) \end{aligned}$$

where $W(w, x, \mu_1, \mu_2)$ is the Lagrangian and $x^*(w)$ is the unique eigenvector corresponding to $\nu_2(w)$. Then, we can use envelope theorem [164] to compute

$$\frac{\partial \nu_2(w)}{\partial w_{ij}} = (x^*(w)_i - x^*(w)_j)^2.$$

Thus, the gradient can be compactly expressed as $\nabla_w \nu_2(w) = p_{x^*(w)}$, where $p_{x^*(w)} := \operatorname{diag}[B^T x^*(w) x^*(w)^T B]\mathbf{1}_{m+n}$, and $\operatorname{diag}[A]$ is the matrix operator that projects all the off diagonal elements to zero and keeps the diagonal untouched. Similarly, we will use $\operatorname{diag}[a]$ to denote the operator that converts the vector a in a diagonal matrix.

When $\nu_2(w)$ is not simple, there are several $x^*(w)$ that solve this optimization

problem. In this case $\nabla_w \nu_2(w)$ is in general not defined, but it is easy to show that for every $x^*(w)$, $p_{x^*(w)} \in \partial_w^+ \nu_2(w)$, where $\partial^+ g(w) := \{p | \langle p, \bar{w} - w \rangle \geq g(\bar{w}) - g(w)\}$ is the concave super-differential of the function $g(w)$. Although in general there is no guarantee of local improvement for every direction $p \in \partial^+ g(w)$, subgradient-type iterations can still reach the global optimum. See [165] for general treatment of subdifferentials of eigenvalues of symmetric matrices.

One interesting consequence of this derivation is that $(p_{x^*(w)})_{ij} \geq 0$. This implies that $\nu_2(w)$ is a nondecreasing function of its elements and the only way to reduce its value is by decreasing some w_{ij}. The main difficulty in our case is that the weights w_{ij} depend on the parameters of the system in a nonlinear manner, i.e. $w_{ij} = |V_i| \left|V_j\right| b_{ij} \cos(\phi_j^* - \phi_i^*)$ where ϕ^* is a solution to

$$F(\phi, b, P) = -Bf(b) \circ \sin(B^T \phi) + P = 0, \tag{4.15}$$

for fixed line inductances b and fixed power schedule P. Therefore, it is not clear at first sight how changes on b and P affect the corresponding w.

We assume that the network is in a stable steady state such that the corresponding ϕ^* is stable and $L(w(\phi^*))$ has only one zero eigenvalue, i.e. $\nu_2(w) > 0$. In the rest of this section we show how changing the network parameters affects $\nu_2(w)$, and how these changes are influenced by the topology of the network and the current operating point.

4.3.1 Power Scheduling

In this subsection we show how power injection changes can locally improve the damping of a power network. We assume fixed line inductances $b = (b_{ij})$

and full control of P within the interior of feasible closed set $B_P = \{P : P_{\min} \leq P \leq P_{\max}\}$. That is, we can not only change the values of P_{g_i}, but we can also change, up to a certain extent, P_{d_i}. Although this used to be an unreasonable assumption, the introduction of renewable energy sources in the distribution part of the network can enable the design of coordination mechanism that produce the desired changes on P_d.

Since b is assumed to be fixed, $F(\phi, b, P) = F(\phi, P)$, and thus every equilibria ϕ^* satisfies, $F(\phi^*, P) = 0$. Here, we will focus on how small changes in the power scheduling $P + \delta P$ affect the position of the equilibrium $\phi^* + \delta\phi$.

Although in principle $\delta\phi, \delta P \in \mathbb{R}^{m+n}$, only a subspace of $\mathbb{R}^{m+n}$ is of interest. Since the network is lossless, $\langle P, \mathbf{1}_{m+n}\rangle = 0$ is always satisfied. Hence we will only consider changes δP in the power schedule s.t. $\delta P \in \ker(\mathbf{1}^T)$. Similarly, since $w(\phi^* + \kappa\mathbf{1}) = w(\phi^*)\ \forall \kappa \in \mathbb{R}$, we will restrict our attention on changes $\delta\phi \in \ker(\mathbf{1}^T)$. The relationship between δP and $\delta\phi$ is then captured by

$$F(\phi^* + \delta\phi, P + \delta P) = F(\delta\phi, \delta P) = 0. \tag{4.16}$$

Theorem 4.2 (Controllability of $\delta\phi$ w.r.t δP)**.** *Given an equilibrium point ϕ^* of* (4.11) *with simple zero eigenvalue, and a power scheduling P satisfying* (4.15)*. There exists a neighborhood of P, $\mathcal{B}_P \subset P + \ker(\mathbf{1}^T)$, and function $\delta\phi(\delta P)$ such that $F(\delta\phi(\delta P), \delta P) = 0, \quad \forall \delta P \in \mathcal{B}_P - P$ and $\delta\phi(\mathcal{B}_P)$ is open relatively to $\ker(\mathbf{1}_{m+n}^T)$, i.e. $\delta\phi$ is fully controllable by δP.*

Moreover, $\frac{d(\delta\phi)}{d(\delta P)}$ can be computed as

$$\frac{d(\delta\phi)}{d(\delta P)} = L(w(\phi^*))^\dagger, \tag{4.17}$$

where $L(w(\phi^*))^\dagger$ is the Moore-Penrose pseudoinverse of the weighted Laplacian $L(w(\phi^*))$.

Proof. The proof of this theorem comes from applying implicit function theorem (see e.g. [121]) on a properly defined function. Notice that $\frac{\partial}{\partial \phi} F(\phi, P) = -L(w(\phi^*))$, and $\frac{\partial}{\partial P} F(\phi, P) = I_{m+n}$. Thus, since $L(w(\phi^*))$ is singular, implicit function theorem cannot be directly applied. However, our restriction of $(\delta\phi, \delta P)$ to the subspace $\ker(\mathbf{1}_{m+n}^T) \times \ker(\mathbf{1}_{m+n}^T)$ does not suffer this problem.

Since both vectors are restricted to $\ker(\mathbf{1}_{m+n}^T)$, by choosing orthonormal basis of column vectors $\{T_j\}$ we can write

$$\delta\phi = Tx \text{ and } \delta P = Ty$$

where the matrix $T = [T_j] \in \mathbb{R}^{(m+n)\times(m+n-1)}$ is a full column rank matrix, $T^T T = I_{m+n-1}$ and $TT^T = I_{m+n} - \frac{1}{m+n}\mathbf{1}_{m+n}\mathbf{1}_{m+n}^T$ is the orthogonal projection onto $\ker(\mathbf{1}_{m+n}^T)$.

Now, define $H(x, y) = T^T F(Tx, Ty)$. Since $\mathbf{1}_{m+n}^T B = 0$, $F(\phi, P) \in \ker(\mathbf{1}_{m+n}^T)$ provided $P \in \ker(\mathbf{1}_{m+n}^T)$. Thus, $F(Tx, Ty) = 0$ if and only if $H(x, y) = 0$, and H represents the same constraints as F when restricted to $\ker(\mathbf{1}^T) \times \ker(\mathbf{1}^T)$. Differentiating H with respect x and y gives

$$\frac{\partial}{\partial x} H(x, y) = -T^T L(w(\phi^*))T, \text{ and } \frac{\partial}{\partial y} H(x, y) = I_{m+n-1}.$$

Since $\frac{\partial}{\partial x} H(x, y)$ is nonsingular, by implicit function theorem, there exist neighborhoods of 0, $\mathcal{B}_x$ and $\mathcal{B}_y$, and a function $x(y)$ such that $H(x(y), y) = 0$ and $x(\mathcal{B}_y) = \mathcal{B}_x$.

Finally, since $H(x(y), y) = 0$ on $\mathcal{B}_y$,

$$\begin{aligned} \frac{d}{dx} H(x(y), y) &= \frac{\partial}{\partial x} H(x(y), P) \frac{dx(y)}{dy} + \frac{\partial}{\partial y} H(x(y), y) \\ &= -T^T L(w(\phi^*))T \frac{dx(y)}{dy} + I_{m+n-1} = 0 \end{aligned}$$

and therefore $\frac{dx}{dy} = (T^T L(w(\phi^*))T)^{-1}$.

Defining $\delta\phi(\delta P) = Tx(T^T\delta P)$ and $\mathcal{B}_P = P + T\mathcal{B}_y$ gives the first statement of the theorem. Equation (4.17) follows from $\frac{d(\delta\phi)}{d(\delta P)} = T\frac{dx}{dy}T^T$, $(T^T L(w(\phi^*))T)^{-1} = T^T L(w(\phi^*))^\dagger T$ and the fact that $TT^T L(w(\phi^*))^\dagger TT^T = L(w(\phi^*))^\dagger$. □

Using (4.17) we can predict how small changes of the power affects the position of the equilibria, which in turn affects the value of $w(\phi)$ (recall $w_{ij} = |V_i|\left|V_j\right| b_{ij}\cos(\phi_j - \phi_i)$). Thus, we can use Theorem 4.2 to compute the changes of the weights δw as

$$\begin{aligned}\delta w = \frac{dw}{d(\delta P)}\delta P &= \left(\frac{\partial w}{\partial \phi}\right)\left(\frac{d(\delta\phi)}{d(\delta P)}\right)\delta P \\ &= \left(-\mathrm{diag}[f(b)\circ\sin(B^T\phi^*)]B^T\right)\left(L^\dagger(w(\phi^*))\right)\delta P \\ &=: A(\phi^*)\delta P\end{aligned}$$

where we use the fact

$$\frac{\partial w}{\partial \phi} = -\mathrm{diag}[f(b)\circ\sin(B^T\phi)]B^T. \tag{4.18}$$

Ideally, we would like to move $\delta w \in \mathrm{span}(p_{x^*}(w(\phi^*)))$, but we are constrained only to the subspace $\mathrm{span}(A(\phi^*))$. Therefore, a natural alternative is to set δP such that the corresponding δw is the orthogonal projection of $p_{x^*(w)}$ onto $\mathrm{span}(A(\phi^*))$. This is done by setting $\delta P = \gamma A(\phi^*)^\dagger p_{x^*(w(\phi^*))}$ which gives

$$\delta w = A(\phi^*)\delta P = \gamma A(\phi^*)A(\phi^*)^\dagger p_{x^*(w(\phi^*))}.$$

Remark 4.3. *Although the updating direction of this section modifies the values of $P_g := (P_{g_i})^T$ and $P_d := (P_{d_i})^T$, it is possible to constraint its actions only to P_g by projecting δP onto $S = \ker(\mathbf{1}_{m+n}^T) \cap B_P$ and setting $P_{d,\min} = P_{d,\max} = P_d$ in the definition of B_P. We denote this projection operation onto the set S by $\Pi_S[\cdot]$, and similarly use $\Pi_{B_P}[\cdot]$ to define the analogous for B_P. Note that $\Pi_{B_P}[\cdot]$ enforces the constraint of B_P.*

4.3.2 Impedance Adaption

In this subsection we study how the changes of line inductances b_{ij}, due to changes in the network topology or the utilization of FACTS devices [166], can affect the operating point of a network. Using this result, we will show that it is possible that the addition of a line can weaken the condition of the network, i.e. $\nu_2(w)$ is reduced.

In order to measure how changes of δb affect the weights w, we proceed in the same manner as in Theorem 4.2. We start by computing the total derivative $\frac{dw}{d(\delta b)}$ which is given by

$$\frac{dw}{d(\delta b)} = \frac{\partial w}{\partial b} + \frac{\partial w}{\partial \phi}\frac{d(\delta \phi)}{d(\delta b)}. \tag{4.19}$$

Since $w(\phi, b) = f(b) \circ \cos(B^T \phi)$, it is straightforward to show that

$$\frac{\partial w}{\partial b} = \operatorname{diag}[v \circ \cos(B^T \phi)] \tag{4.20}$$

where the elements of the vector v are $(v)_{ij} = |V_i| \left|V_j\right|$ if $ij \in \mathcal{E}$ and $(v)_{ij} = 0$ otherwise.

The main difficulty again rises in computing how the changes of b, i.e. δb, affect ϕ^*. This is assessed in the next theorem. As in Theorem 4.2, we restrict our attention to $\delta\phi \in \ker(\mathbf{1}^T)$. We do not impose any restriction on δb besides the physical ones, i.e. $b + \delta b \in \{b : b_{\min} \leq b \leq b_{\max}\}$.

Theorem 4.3 (Controllability of $\delta\phi$ w.r.t δb)**.** *Given an equilibrium point ϕ^* of* (4.11) *with a simple zero eigenvalue, and bus admittances b such that $F(\phi^*, b) = 0$. There exists a neighborhood of b, $\mathcal{B}_b$, a neighborhood of ϕ^*, $\mathcal{B}_{\phi^*}$, and function $\delta\phi(\delta b)$ such that*

$$F(\delta\phi(\delta b), \delta b) = 0, \quad \forall \delta b \in \mathcal{B}_b - b$$

and $\phi^* + \delta\phi(b + \mathcal{B}_b) = \mathcal{B}_{\phi^*}$. *Moreover,* $\frac{d(\delta\phi)}{d(\delta P)}$ *can be computed as*

$$\frac{d(\delta\phi)}{d(\delta b)} = -L(w(\phi^*))^\dagger B\mathrm{diag}[v \circ \sin(B^T\phi^*)]. \tag{4.21}$$

Proof. Since $\delta\phi$ is restricted to $\ker(\mathbf{1}^T)$ we can use the same transformation T to transform $F(\delta\phi, \delta b) = 0$ into

$$H(x, \delta b) = T^T F(Tx, \delta b) = 0.$$

The Jacobian $\frac{\partial}{\partial x}H(x, \delta b) = -T^T L(w(\phi^*))T$ remains the same, and

$$\frac{\partial}{\partial b}H(x, \delta b) = -T^T B\mathrm{diag}[v \circ \sin(B^T\phi)].$$

Therefore, since $\frac{\partial}{\partial x}H(x, \delta b)$ is nonsingular, we can apply again implicit function theorem to get

$$\frac{dx(\delta b)}{d(\delta b)} = -\left(\frac{\partial}{\partial x}H(x, \delta b)\right)^{-1} \frac{\partial}{\partial b}H(x, \delta b).$$

Equation (4.21) follows after reverting the change of variables.

□

Now substituting (4.18), (4.20) and (4.21) into (4.19) we obtain

$$\begin{aligned}\frac{dw}{db} &= \mathrm{diag}[v \circ \cos(B^T\phi^*)] \\ &\quad + \mathrm{diag}[f(b) \circ \sin(B^T\phi)](R)\mathrm{diag}[v \circ \sin(B^T\phi^*)]\end{aligned}$$

where $R = B^T L(w)^\dagger B$ is the effective resistance matrix when the weights w are interpreted as conductances. $R_{ij,kl}$ represents the voltage difference between nodes i and j when a current of 1 unit is injected in i and subtracted from j [167]. Notice that in our case, it is possible that some of the weights w_{ij} are negative. Nonetheless $L(w)$ is positive semi-definite with a single zero eigenvalue and

therefore $R_{ijij} > 0$. Thus, we can still interpret R_{ijij} as a measure of the distance between i and j.

Clearly, using this notion of distance, one can see how $\nu_2(w)$ is more sensitive to changes between nodes "farther" away. However, what is interesting here is the appearance of the term $\cos(\phi_j^* - \phi_i^*)$. When the phase difference between certain buses is larger than $\frac{\pi}{2}$ we have $\cos(\phi_j^* - \phi_i^*) < 0$ and therefore and increment on b_{ij} could possibly affect negatively the weight $w_{ij}(\phi)$. This phenomenon is numerically illustrated in section 4.5.1.

4.4 Dynamics-aware Optimal Power Flow

In this section we show that the use of pseudospectral abscissa $\alpha_\varepsilon(A)$ provides a convenient framework that not only balances transient amplitude and asymptotic convergence rate, but also jointly guarantees voltage and small signal stability. This subsequently leads to a new optimization formulation that can jointly enforce both stability constraints with a single performance metric.

Given $\varepsilon \geq 0$ the pseudospectrum Λ_ε of a matrix A is defined as the set of eigenvalues of all matrices $X \in \mathbb{C}^{n\times n}$ satisfying $||X - A||_2 \leq \varepsilon$ where $||\cdot||_2$ is the spectral norm. With this notation, the pseudospectral abscissa is defined by

$$\alpha_\varepsilon(A) = \max\{\Re[z] : z \in \Lambda_\varepsilon(A)\}.$$

When $\varepsilon = 0$, $\alpha_0(A)$ reduces to the spectral abscissa which is equivalent to the constraint in [112]. There are several advantages on using pseudospectral abscissa instead, which we now summarize.

- Unlike $\alpha_0(A)$, $\alpha_\varepsilon(A)$ is locally Lipschitz with respect to A and thus easier to numerically compute.
- If $\beta(A)$ is the distance to instability, then the following relationship follows:

$$\alpha_\varepsilon(A) \leq 0 \iff \beta(A) \geq \varepsilon \iff \mathcal{H}_\infty(A) \leq \frac{1}{\varepsilon}$$

  Here, $\mathcal{H}_\infty(A)$ is the $\mathcal{H}_\infty$ norm of the system [168], i.e. $\mathcal{H}_\infty(A) = \sup_{\omega\in\mathbb{R}} \sigma_{\max}(H(j\omega))$, where $\sigma_{\max}(H(s))$ is the maximum singular value of the transfer function $H(s) = (A - sI)^{-1}$.
- $\alpha_\varepsilon(A)$ captures several dynamic properties for different values of ε. For $\varepsilon = 0$, $\alpha_\varepsilon(A)$ is the asymptotic rate. If $\alpha_\varepsilon(A) = 0$ then $\varepsilon^{-1} = \mathcal{H}_\infty(A)$ and when $\varepsilon \to +\infty$, $(\alpha_\varepsilon(A) - \varepsilon)$ is the initial rate of decay [169].

With these nice properties, we now propose the following optimization problems to study the performance limits of a power network.

$$\mathbf{H}_\infty: \quad \underset{\varepsilon\geq 0, x, z, u}{\text{minimize}} \quad h(\varepsilon)$$

$$\text{subject to} \quad (4.4) \tag{4.22a}$$

$$\alpha_\varepsilon(A(x, z, u, v)) \leq 0 \tag{4.22b}$$

Remark 4.1 guarantees that by satisfying (4.4a) we can find (x, z, u, v) that satisfies the equilibrium equations of (4.6a)-(4.6b) and therefore we do not need (4.6a)-(4.6b) as constraints.

The function $h(\varepsilon)$ is decreasing, which guarantees that the optimal solution $(\varepsilon^*, x^*, z^*, u^*)$ of $\mathbf{H}_\infty$ has the constraint (4.22b) met with equality and makes $\frac{1}{\varepsilon^*} = \mathcal{H}_\infty(A(x^*, z^*, u^*, v))$. Thus, this problem finds the optimal configuration in terms of $\mathcal{H}_\infty(A(x, z, u, v))$. Furthermore, the solution of $\mathbf{H}_\infty$ also guarantees voltage stability, since (4.6a)-(4.6b) has a stable equilibrium, and ensures a robust

stability radius of ε^*. Here we will use $h(\varepsilon) = -20\log_{10}(\varepsilon)$, which amounts to the maximum power gain in decibels (dB) of the transfer function $H(s)$ when $\alpha_\varepsilon(A(x,z,u,v)) = 0$.

Alternatively, one could choose to sacrifice $\mathcal{H}_\infty$ optimality by minimizing $\alpha_\varepsilon(\cdot)$ for fixed ε. That is,

$$\begin{aligned} \mathbf{A}_\varepsilon: \quad & \underset{x,z,u}{\text{minimize}} \qquad \alpha_\varepsilon(A(x,z,u,v)) \\ & \text{subject to} \qquad (4.4) \end{aligned} \tag{4.23}$$

When $\varepsilon = 0$, $\mathbf{A}_\varepsilon$ finds the optimal configuration u^* that has the fastest asymptotic rate. On the other hand, when $\varepsilon \to +\infty$ the solution of the problem provides a u^* that optimizes the *initial decay rate* of a small perturbation [169].

This new formulation also unveils a fundamental tradeoff between voltage stability and small signal stability of power networks that has not been previously analyzed. Finding the maximum distance to voltage collapse implies using $\alpha_0(A)$ in $\mathbf{A}_\varepsilon$. While the solution of this problem will be optimal in terms of voltage stability margin, it can potentially have transients with large amplitude. On the other hand, if one is interested in minimizing the $\mathcal{H}_\infty(A)$ using $\mathbf{H}_\infty$, then the required voltage stability margins might not be met.

The optimization problems $\mathbf{H}_\infty$ and $\mathbf{A}_\varepsilon$ conform a novel framework that can be readily combined with the **OPF**. They provide a unifying representation of several dynamical properties within a one parameter family of functions $\alpha_\varepsilon(\cdot)$. This is very convenient as the operator can choose different values of ε, depending on the different needs of the power network in consideration.

This results in the following formulation for a Dynamics-aware Optimal

Power Flow problem.

$$\textbf{Dyn-OPF}: \quad \underset{\varepsilon \geq 0, x, z, \mu}{\text{minimize}} \quad c(V, P_G, Q_G) \tag{4.24}$$

$$\text{subject to} \qquad (4.4)$$

$$h(\varepsilon) \leq h^* \tag{4.25}$$

$$\alpha_\varepsilon(A) = 0 \tag{4.26}$$

$$\alpha_{\hat{\varepsilon}}(A) \leq a^* \tag{4.27}$$

where $\hat{\varepsilon}$ is a constant parameter.

The solution to **Dyn-OPF** will provide an operating point that minimizes the generation cost and keeps a maximum power gain of $20\log_{10}(\mathcal{H}_\infty(A)) \leq h^*$. On the other hand, it is possible to use (4.27) to provide additional constraints on the system. For example, by setting $\hat{\varepsilon} = 0$, (4.27) can be used to impose specific voltage stability margins. Notice that since neither the **OPF** nor $\mathbf{H}_\infty$ and $\mathbf{A}_\varepsilon$ are convex problems, all their solutions and the solution to **Dyn-OPF** are local minima.

4.5 Numerical Examples

This section provides numerical examples that illustrate the findings of this chapter.

4.5.1 Improving the Damping

We start by showing the effect of using the updating directions described in section 4.3 on two simple network configurations.

Generator Power Scheduling

Consider a network of 3 generator buses, and 3 load buses disposed in a complete graph configuration as in Figure 4.2 with $b_{ii} = 10$ for every generator, $b_{12} = b_{13} = 2$ and $b_{23} = 10$. The initial power scheduling is

$$P_d = [\ 4 \quad 6 \quad 8\ ]^T, \text{ and } P_g = [\ 7.994 \quad 3.006 \quad 7\ ]^T.$$

Assume also that the power demanded in each bus is fixed and cannot be modified by the algorithm.

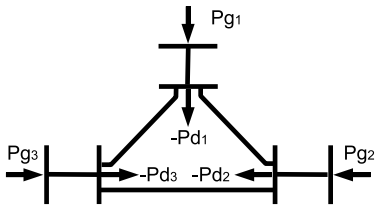


Figure 4.2: 3 Bus Power Network

We now see how the operating point of the network can be locally improved. One possible equilibrium ϕ^* that solves (4.16), for given P, is

$$\phi^* = [\ .513 \quad 0 \quad .032 \quad .808 \quad .097 \quad .279\ ]^T \pi.$$

The values of P chosen are such that the system is very close to the bifurcation. Figure 4.3 shows the evolution of $\Re[\lambda_2]$, the corresponding upper bound provided in Theorem 4.1 and the location of the 4th closest eigenvalues to the imaginary axis. We can see that when the system is close to a bifurcation, not only our proposed adaptation is more effective, but also the upper bound computed is tighter.

Initially $\Re[\lambda_2] = \lambda_2$ is very close to zero, but it gradually decreases until a new eigenvalue with non zero imaginary part becomes dominant. This is

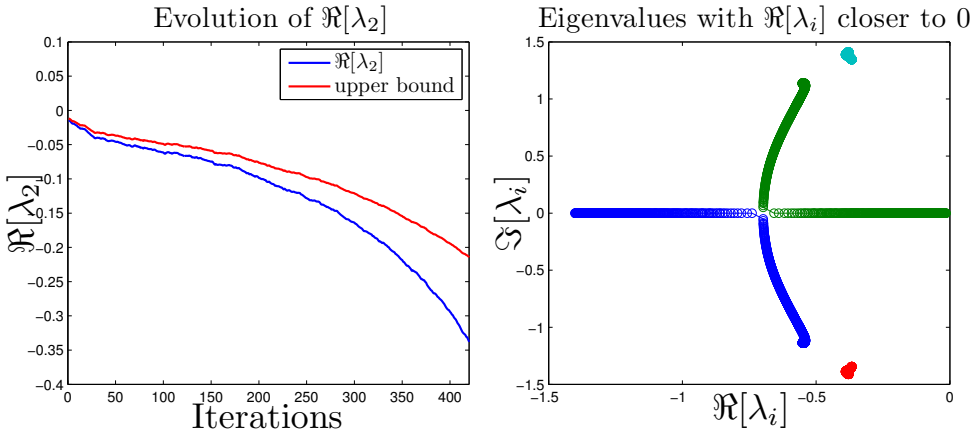


Figure 4.3: Evolution of $\Re[\lambda_2]$

captured in the right graph of Figure 4.3. After this point, $\Re[\lambda_2] = -\frac{d(x_i)}{2m(x_i)}$ and the dependence on the algebraic connectivity is lost.

Adding or Removing a Line

In this example each generator bus g_i is generating $P_{g_i} = \bar{P}$ and each load bus demands a power of $P_{d_i} = \bar{P}$ with $\bar{P} = 5$. We assume $b_{ij} = 10\ \forall ij$.

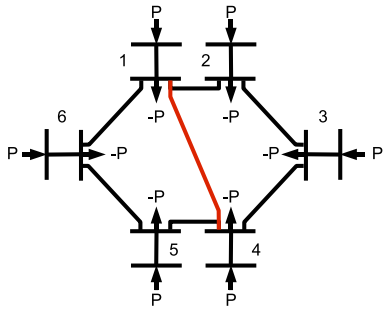


Figure 4.4: 6 Bus Power Network

Among the possible equilibria for this network we study the equilibrium ϕ_2^* given by $(\phi_2^*)_{d_i} = \frac{2\pi}{6}(i-1)$, $i \in \{1,\ldots,6\}$ and $(\phi_2^*)_{g_i} = (\phi_2^*)_{d_i} + \arcsin(\frac{1}{2})$.

In this case, when we add a line between d_1 and d_4, i.e. we increase $\sigma b_{d_1 d_4}$ (in red) from $\sigma = 0$ to $\sigma = 1$, $\frac{dw}{db}$ becomes

$$\frac{dw}{db_{d_1 d_4}} = -e_{d_1 d_4} < 0,$$

since $\phi^*_{d_4} - \phi^*_{d_1} = \pi$, and therefore $\langle \nabla \nu_2, \delta w \rangle \leq 0$. In fact, since $x^*(w(\phi^*_2))_{d1} \neq x^*(w(\phi^*_2))_{d4}$, we can see in Figure 4.5 that $\nu_2(w(\phi^*_2))$ decreases.

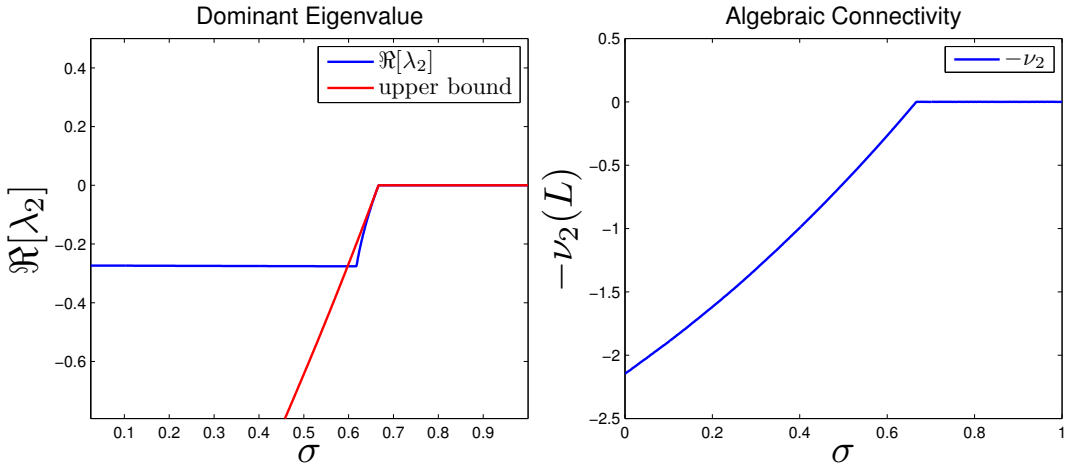


Figure 4.5: Effect of Adding $b_{d_1 d_4}$

Here we can also see that although $\Re[\lambda_2]$ and $\nu_2(L(w(\phi^*)))$ might not be in general correlated, when the system is close to a bifurcation, i.e. $\nu_2(L(w(\phi^*)))$ and $\Re[\lambda_2]$ are close to zero, the changes on $\nu_2(L(w(\phi^*)))$ directly affect $\Re[\lambda_2]$. Moreover, Figure 4.5 shows how our upper-bound is only valid in this specific case.

4.5.2 Test Cases Dynamics-aware OPF

Finally, we provide two examples to illustrate the properties of the optimization framework presented in section 4.4. The dynamic models of (4.6a) and (4.9) as

well as the algebraic equations (4.4a) and (4.6b) are computed using the *Power System Toolbox* (PST) [170]. $\alpha_\varepsilon(A)$ is evaluated using the Matlab code provided with [171] with a tolerance of $1e-12$. The gradients of $\alpha_\varepsilon(A)$ are computed numerically and the Matlab Optimization Toolbox is used to compute the local optimum. We call the *fmincon* subroutine with function and constraint tolerance of $1e-6$ for optimizations involving $\alpha_\varepsilon(A)$ ($\mathbf{H}_\infty$, **Dyn-OPF** , $\mathbf{A}_\varepsilon$) and with tolerance $1e-7$ for **OPF** . All the results presented in this section are in base 100MVA.

The cost function $c(V, P_G, Q_G)$ used is the standard quadratic cost function depending only on the active generation, i.e.

$$c(P_G) = \sum_{i \in \mathcal{N}} c_{2i} P_{G_i}^2 + c_{1i} P_{G_i} + c_{0i}.$$

This framework is not limited to this specific $c(V, P_G, Q_G)$ and can be easily extended to consider other objective functions.

Two Area Test Case

This example illustrates properties and differences between the local minima of the optimization problems $\mathbf{H}_\infty$, $\mathbf{A}_\varepsilon$ and **OPF**. We consider a 2 area power network with 13 buses and 4 generators with detailed 2-axis subtransient generators, static exciters, power system stabilizers and 2 induction motors on the load buses 4 and 14.

The load profile as well as the parameters of the induction motors were take from the file `d2asbegp.m` that comes with the PST distribution. The generator dynamics parameters are chosen homogeneously and listed in Table 4.1.

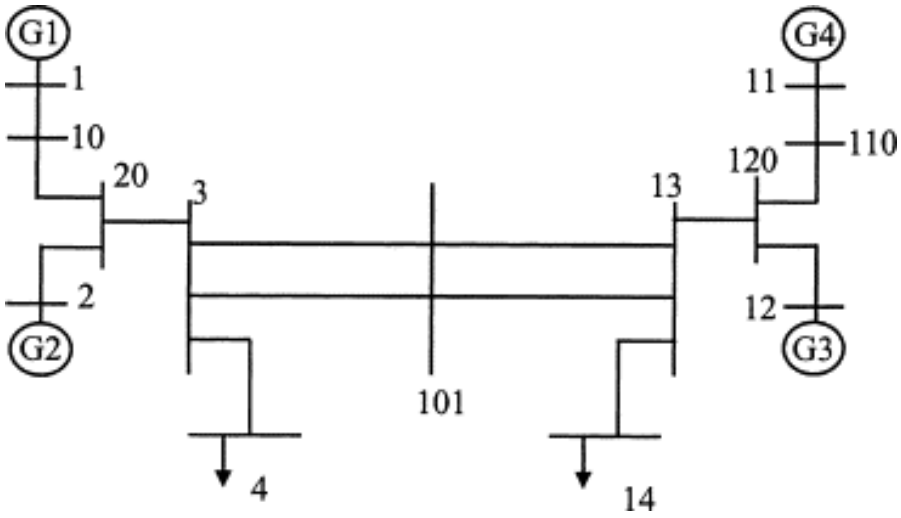


Figure 4.6: Two area 13-bus test case

Table 4.1: Generator dynamics parameters for the two are test case

Gen #	x_l (pu)	r_a (pu)	x_d (pu)	x'_d (pu)
1,2,3,4	0.022	0	0.2	0.033
x''_d(**pu**)	T'_{do}(**sec**)	T''_{do} (sec)	x_q (pu)	x'_q (pu)
0.028	8	0.03	0.189	0.061
x''_q (pu)	T'_{qo} (sec)	T''_{qo} (sec)	**H** (sec)	$d_o = d_1$ (pu)
0.027	0.4	0.05	58.5	0

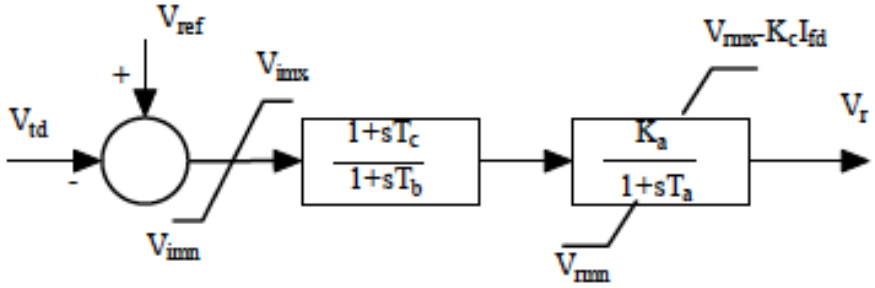


Figure 4.7: AC4a Excitation System

Generators are provided with identical AC4a excitation systems and PSSs. Figure 4.7 shows a block diagram of the AC4a system, where T_r is the transducer time constant, K_a and T_a are the voltage regulator gain and time constants, respectively, $[V_{i_{\min}}, V_{i_{\max}}]$ are the input voltage saturation limits, $[V_{r_{\min}}, V_{r_{\max}}]$ are the output voltage saturation limits (we take $K_c = 0$ in Figure 4.7) and T_b and T_c are compensator constants.

Table 4.2: AC4a excitation system parameters

Gen #	$\mathbf{T_r}$ (sec)	$\mathbf{K_a}$	$\mathbf{T_a}$ (sec)	$\mathbf{T_b}$ (sec)
1,2,3,4	.0145	200	.05	0
$\mathbf{T_c}$ (sec)	$\mathbf{V_{i_{min}}}$ (pu)	$\mathbf{V_{i_{max}}}$ (pu)	$\mathbf{V_{r_{min}}}$ (pu)	$\mathbf{V_{r_{max}}}$ (pu)
0	-10	10	-4.53	5.64

We use standard PSSs with washout filter and two lag compensators with Laplace Transfer

$$H_i^{PSS}(s) = \kappa_i \frac{sT_{w,i}}{1+sT_{w,i}} \frac{1+sT_{n1,i}}{1+sT_{d1,i}} \frac{1+sT_{n2,i}}{1+sT_{d2,i}}$$

with equal parameters $\kappa = 1$, $T_w = 10$, $T_{n1} = .05$, $T_{d1} = .02$, $T_{n2} = .08$ and $T_{d2} = .015$. All time constants are in seconds.

We solve **OPF**, $\mathbf{H}_\infty$ and $\mathbf{A}_\varepsilon$ with $\varepsilon = 0$. We assume equal cost among the four generators with parameters $c_0 = 0$ and $c_1 = c_2 = 1$. The optimal power scheduling is illustrated in Table 4.3. Table 4.4 shows the asymptotic rate of convergence $\alpha_0(A)$, minimum damping ratio ξ and maximum gain of the system $\mathcal{H}_\infty(A)$ for the three optimization problems studied in this test case, and Figure 4.8 shows the corresponding critical eigenvalues.

This example clearly illustrates the tradeoff between asymptotic rate of con-

Table 4.3: Power Scheduling of two area 13-bus test case for $\mathbf{H}_\infty$, **OPF** and $\mathbf{A}_\varepsilon$ with $\varepsilon = 0$

Gen #	$\mathbf{H}_\infty$		**OPF**		$\mathbf{A}_\varepsilon$	
	P_G	Q_G	P_G	Q_G	P_G	Q_G
1	6.64	1.04	4.90	0.86	5.86	2.33
2	7.81	2.12	5.01	0.02	5.69	1.65
3	3.59	-1.66	4.89	0.87	5.32	0.72
4	2.00	1.23	5.01	-1.13	3.11	1.51

Table 4.4: Dynamic performance metrics of different operating solutions

	$\mathbf{H}_\infty$	**OPF**	$\mathbf{A}_{\varepsilon=0}$
$\alpha_0(A)$	-0.100238	-0.100331	-0.100598
$\xi(A)$	0.1076	0.0571	0.0108
$\mathcal{H}_\infty(A)$ (dB)	38.23	40.60	55.75

vergence and oscillations amplitude. In particular, we can see how $\mathbf{A}_{\varepsilon=0}$ succeeds in obtaining a smaller $\alpha_0(A)$ than $\mathbf{H}_\infty$ and **OPF** , but it performs very poorly in terms of minimum damping ratio ξ and maximum frequency gain $\mathcal{H}_\infty(A)$. This confirms our claim suggesting that $\alpha_0(A)$ should not be used as a performance metric in order to avoid oscillations like in [112] as it can potentially amplify them. This is somehow counterintuitive since $\alpha_0(A)$ does succeed in avoiding Hopf Bifurcations.

On the other hand $\mathbf{H}_\infty$ clearly outperforms **OPF** in damping the modes achieving a relative increment in the minimum damping of $\frac{\xi_{\mathbf{H}_\infty}}{\xi_{\mathbf{OPF}}} = 2.83$, almost three times higher, and a gain reduction $\frac{\mathcal{H}_{\mathbf{H}_\infty}(A)}{\mathcal{H}_{\mathbf{OPF}}(A)} = -2.37$ dB. Thus, this example also shows how the dynamic behavior of a power network can be considerably

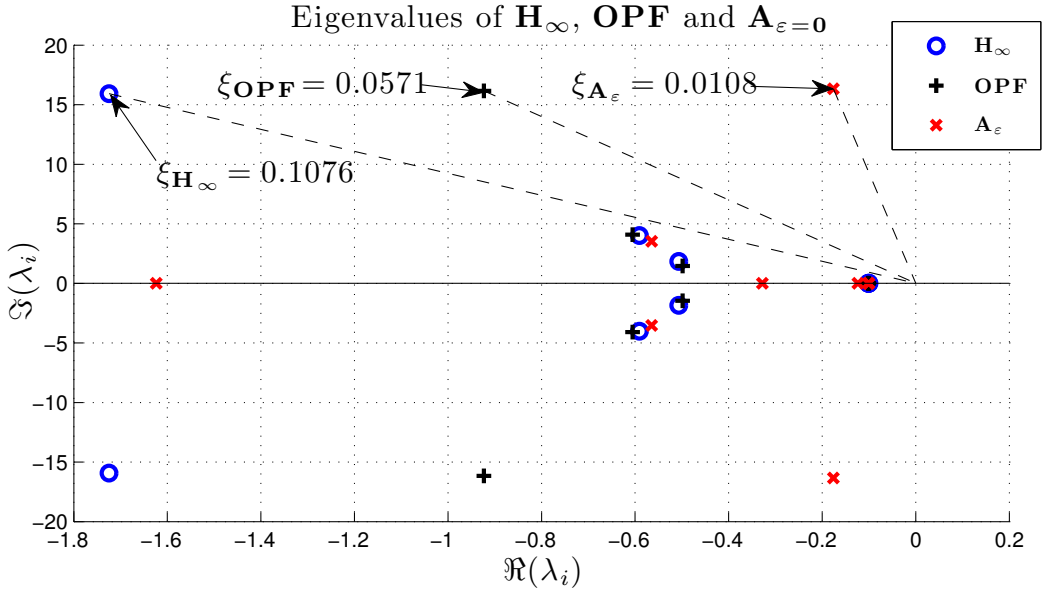


Figure 4.8: Eigenvalues of the two are test system in Firgure 4.6 for the output of **OPF**, $\mathbf{A}_\varepsilon$ with $\varepsilon = 0$ and $\mathbf{H}_\infty$. The counter-clockwise angle between the dashed lines and the horizontal axis θ defines the damping ratio ($\xi = \cos(\theta)$) . Only the eigenvalues closer to the imaginary axis are shown.

improved by solely changing the operating point.

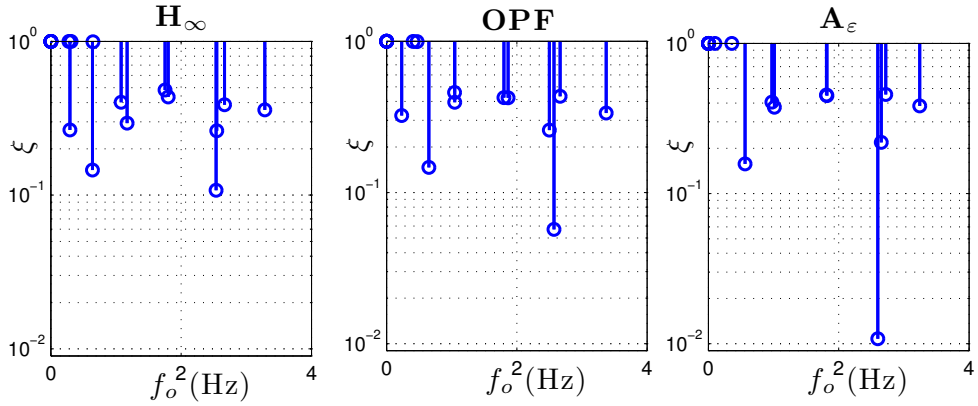


Figure 4.9: Modes vs frequency of the two are test system solutions to $\mathbf{A}_\varepsilon$, **OPF** and $\mathbf{H}_\infty$.

Finally, we present in Figure 4.9 a stem graph of the system modes (damping vs frequency) for the different operating points computed. It is interesting to

notice that some modes do not change considerably by modifying the power scheduling. This evidences the limits of the framework. That is, if the mode that defines the minimum damping is not very sensitive to the power scheduling, then the improvement may not be considerable.

Therefore, while this method is effective to alleviate possibly stressed scenarios cause by a poor scheduling, it is certainly not a substitute to current industry practices of controller designs which are clearly needed to modify the modes that are not sensitive to the scheduling.

New England Power Grid

We now consider the IEEE 39-bus New England power grid with 10 detailed 2-axis generator models shown in Figure 4.10. Generators 1 to 9 are equipped with AC4a excitation system with parameters described also by Table 4.2 and PSSs using the optimal configuration described in [71]. The dynamic data of the generators was obtained from [172]. We select generator 10 as infinite bus in order to eliminate the zero eigenvalue of the system.

In order to illustrate a stressed state of the network, we define two different generation cost values. Generators 1, 8-10 use parameters $c_2 = 0.01$, $c_1 = 3.0$ and $c_0 = 0.0$, and generators 2-7 use $c_2 = 0.01$, $c_1 = 0.3$ and $c_0 = 0.0$. This creates a power transfer from area 2 to area 1 of Figure 4.10 through lines (15.17), $(3,4)$ and $(9,39)$ and thus brings the system closer to its stability boundary.

We first solve the **OPF** and $\mathbf{H}_\infty$ problems with voltage constraints limits of $[0.9, 1.1]$ (pu) for every load bus and $[0.95, 1.05]$ (pu) for every generator bus. Generation limits are set homogeneously to $P_{G_i}^{\max} = 11$, $P_{G_i}^{\min} = 0$, $Q_{G_i}^{\max} = 8$

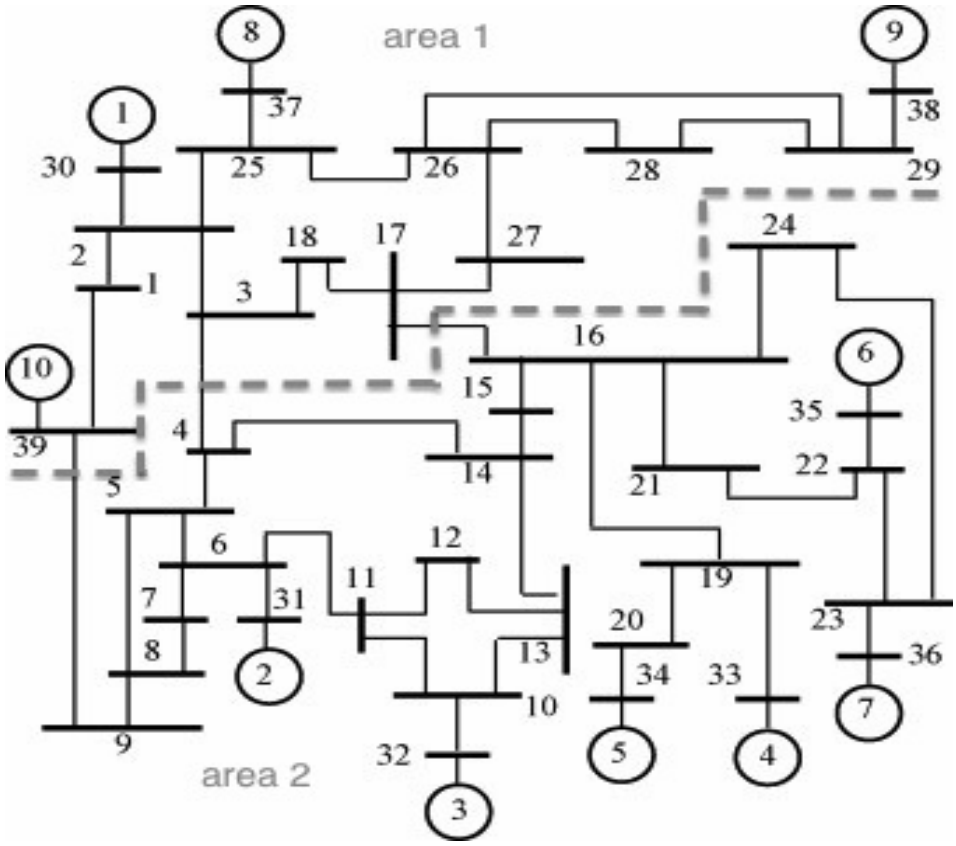


Figure 4.10: One line diagram of New England 39-bus system

and $Q_{G_i}^{\min} = -5$. All flow and thermal constraints are made non-binding. The solution of $\mathbf{H}_\infty$ gives a value of $h(\varepsilon^*) = 32.392$ dB while for the optimum of **OPF** $h(\varepsilon^*) = 32.808$ dB. The relative damping ratio gain is $\frac{\xi_{\mathbf{H}_\infty}}{\xi_{\mathbf{OPF}}} = 2.71$ which indicates a significant increment on the system damping.

However, this damping improvement implies an increase of the generation cost from $c(P_G^*) = 59.4$ in **OPF** to $c(P_G^*) = 112.5$ which amounts to a 112.0% increment. This is quite inefficient and we would like to balance the tradeoff between economic efficiency and dynamics performance. We therefore run our **Dyn-OPF** using $h^* = 32.398 \in [32.392, 32.808]$ dB and $a^* = 0$.

Figure 4.11 shows the modes stem graphs for the three different optimization

Table 4.5: Power Scheduling of **OPF** , $\mathbf{H}_\infty$ and **Dyn-OPF** with $h^* = 32.398$ and $a^* = 0$

Gen #	**OPF**		$\mathbf{H}_\infty$		**Dyn-OPF**	
	P_G	Q_G	P_G	Q_G	P_G	Q_G
1	0.00	1.64	1.97	2.19	10.75	1.84
2	7.75	4.77	10.93	5.01	10.98	4.93
3	7.53	6.78	4.64	5.75	5.47	5.72
4	9.55	5.26	2.40	3.37	2.00	3.38
5	9.09	3.48	10.98	3.16	10.99	3.12
6	10.53	5.33	0.18	2.32	1.34	2.64
7	7.73	2.42	0.71	0.45	8.92	1.63
8	0.00	1.82	11.00	1.56	0.94	1.82
9	0.66	1.09	8.61	0.75	0.01	0.76
10	9.95	2.89	11.00	2.72	11.00	2.29

problems solved as well as the generation cost incurred by each. We can see that by allowing a generation cost of $c(P_G^*) = 86.0$, i.e. a 61.9% increment, we are able to obtain a damping ratio gain of $\frac{\xi_{\textbf{Dyn-OPF}}}{\xi_{\textbf{OPF}}} = 2.02$. The corresponding eigenvalues are shown in Figure 4.12. Although this cost increment might be unfeasible for regular operation, it can certainly be afforded in order to momentarily avoid an unexpected stressed condition.

The frequency that maximizes $\mathcal{H}_\infty(A)$ is $\omega = 0$. A detailed analysis of the left and right singular vectors of the singular value $\sigma_{\min}(A) = \sigma_{\max}(j0I - A)^{-1} = \mathcal{H}_\infty(A)$ for the solutions of $\mathbf{H}_\infty$ and **OPF** shows that the high gain of the system transfer function $H(s) = (sI - A)^{-1}$ is achieved between PSS state variables of several groups of generators. This suggests that the system configuration is in a

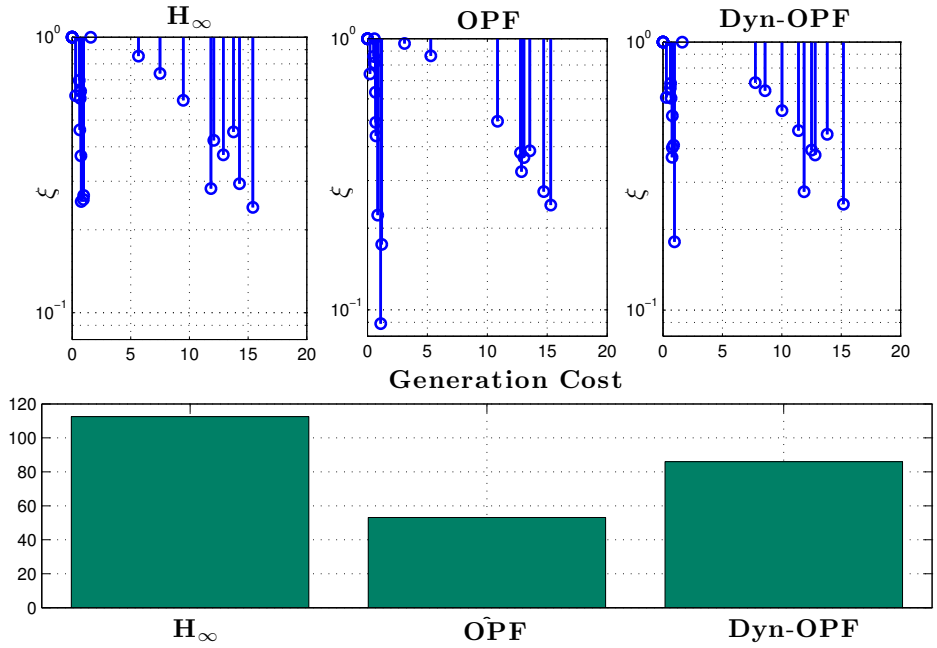


Figure 4.11: Damping ratios and generation cost of New England power grid

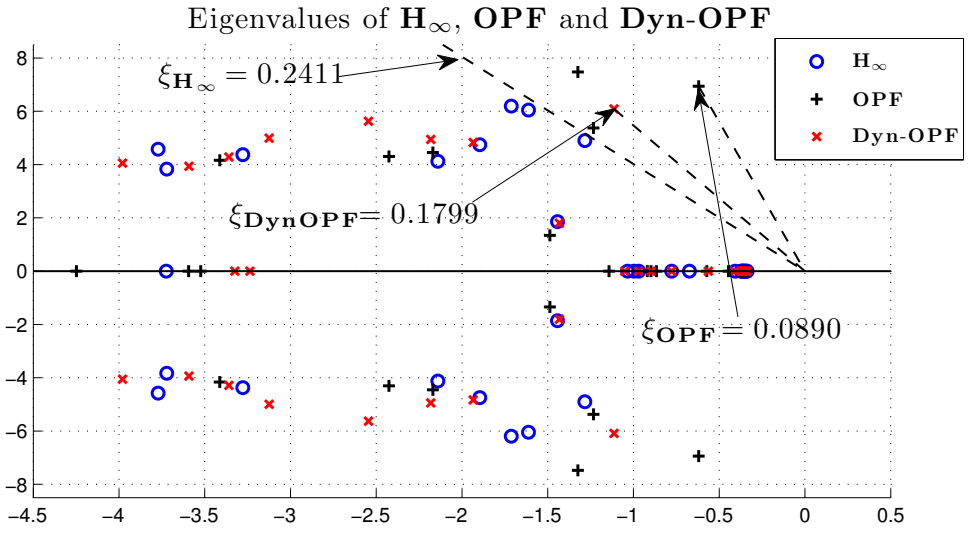


Figure 4.12: Critical eigenvalues of New England power grid. The counter-clockwise angle between the dashed lines and the horizontal axis θ defines the damping ratio ($\xi = \cos(\theta)$)

point that is mostly sensitive to changes on the PSSs parameters. It also explains the differences between the power schedulings on Table 4.5 and the little gain reduction of -0.41 dB from **OPF** to $\mathbf{H}_\infty$, i.e. one needs considerable changes on the scheduling in order to slightly improve $\mathcal{H}_\infty(A)$.

Furthermore, since the operating point has changed with respect to the one used in [71] to compute the optimal PSS parameters, a different configuration could further reduce the damping of the system. This can be easily included in our framework and is subject of current research.

CHAPTER 5

FUTURE WORK

This thesis covers several aspects of collective synchronization on networks that spans from theoretical guarantees and performance analysis to protocol implementation and parameters optimization. Besides the specific extensions that each of the individual lines of work of this thesis might have, there are some general directions that are of general interest.

One example concerns the interplay between network topology and system performance. In chapters 3 and 4 we have seen that the topology of the network as well as its parameters have a direct effect on the performance of the system. We saw in both cases that by changing the value of the graph weights as well as its topology we can obtain significant changes on its performance. In particular, the effect of noise in the system is affected by the topology. In other words, the agents can use the network to collectively reduce the noise in the system and significantly outperform individuals. Understanding this relationship is of great practical interest and can be used to improve the performance of distributed systems. Some related work is present in [173, 174].

Another interesting direction worth pursuing is related to the notion of convergent measure discussed in remark 3.2. The notion of invariant measure, that appeared in section 2.4.1 to compute the final frequency the system converges, is widely used in consensus systems [175–177] to compute the consensus values, see e.g. [178]. The results presented on chapter 3 suggest that it is possible to design a system that posses time varying measures whose behavior can be controlled. Thus, by using local interactions within a network it is possible to control and probably optimize global measures of the whole system.

We also present several directions that we consider worth pursuing on each of the chapters of this thesis.

5.1 Coupled Oscillators

For example, in homogeneous coupled oscillators, we are interested in further understanding how the value of b needed for in-phase synchronization depends on the topology. Besides completing the proof in section 2.2.3 for the cases when $m = 1, 3, 5$, it would be interesting to see whether b can be bounded by a measure of the connectivity of the graph. Our intuition tells us that such result should be provable, yet we have not be able to obtain it.

Additionally, we are interested in eliminating the isolated orbit assumption of section 2.4.2 and see whether the convergence analysis can be also extended to the alternative solution discussed in remark 2.6. The main advantage of this alternative solution is that it only uses phase difference information. Although this was a disadvantage in chapter 3 as it produced backward jumps, it makes it quite suitable to implement using pulse-coupling. Therefore, it can be used to improve the performance of recent protocols that are based on models of homogeneous frequency pulse-copuled oscillators which are unavoidably implemented using heterogeneous frequencies, see e.g. [179].

5.2 Skewless Network Clock Synchronization

Besides showing that skew information is not needed to synchronize the clocks of networked nodes, perhaps the most striking result of chapter 3 is the fact

that in the presence of noise one cannot escape from having a leader or orchestrator when one seeks to achieve consensus in time and frequency at the same time. This is not a problem in our application itself as we usually have a specific source of time (UTC) that we seek to follow.

However, it does seem to become restrictive in more distributed applications such as sensor networks, where one only needs a common relative notion of time. The main difficulty is that the same property that is used to guarantee consensus, i.e. the zero eigenvalue on the Laplacian matrix, is the one that allows the noise to accumulate and drift the system away. It would be interesting to design a control law that is robust to this issue.

Additionally, we are interested in investigating the possibility of decentralizing the parameter optimization. So far, we have been able to use numerical methods to find locally optimal parameters. These methods are centralized and therefore unfitted for distributed applications.

5.3 Dynamics-aware OPF

Besides the insightful results presented on sections 4.2 and 4.3, we consider that the most promising direction to pursue is the one related with our dynamics-aware optimization framework. Although in section 4.4 we have shown that is possible to integrate several dynamic performance metrics to the OPF problem, there is still a gap that prevents its application in real systems. Our current formulation solves the problem using gradient-based methods that can perform very poorly. As future work we are interested in investigating the development of more efficient numerical methods that are able to handle several thousands

of variables.

We are also interested in expanding our framework to include additional performance metrics such as $\mathcal{H}_2$ norm and to include controller synthesis. Optimal controllers are usually designed based on a fixed base operating point. However, as the state of the grid changes the designed controllers are no longer optimum. In order to cope with the future challenges of the incursion of renewable generation, the future grid must be able to adapt and reconfigure the controlling scheme online.

APPENDIX A

APPENDIX

A.1 Proof of Theorem 2.3

Proof. As in Theorem 2.2 we will use our cut condition to show the instability of ϕ^*. Thus, we define a partition $P = (S, V(G)\backslash S)$ of $V(G)$ by taking S to be a maximal subset of $V(G)$ such that $d(\phi, S) < \frac{4\pi}{m}$, see Figure 2.9 for an illustration of P. Notice that any of these partitions will include all the oscillators of two consecutive blocks of every constellation.

Instead of evaluating the total sum of the weights in the cut we will show that the sum of edge weights of the links connecting the nodes of one constellation in S with the nodes of a possibly different constellation in $V(G)\backslash S$ is negative. In other words, we will focus on showing

$$\sum_{ij \in \mathcal{K}_{l_1 l_2}} f'(\phi_j^* - \phi_i^*) < 0 \tag{A.1}$$

where $\mathcal{K}_{l_1 l_2} = \{ij : i \in C_{l_1} \cap S, j \in C_{l_2} \cap V(G)\backslash S\}$.

Given any subset of integers J, we define

$$g_m^J(\delta) = g_m(\delta) - \sum_{j \in J} f(\frac{2\pi}{m} j + \delta).$$

Then, we can rewrite (A.1) as

$$\sum_{ij \in \mathcal{K}_{l_1 l_2}} f'(\phi_j^* - \phi_i^*) =$$

$$\begin{aligned}
&= (g_m^{\{0,1\}})'(\delta_{l_1 l_2}) + (g_m^{\{-1,0\}})'(\delta_{l_1 l_2}) \\
&= 2g_m'(\delta_{l_1 l_2}) - f'(\delta_{l_1 l_2} + \frac{2\pi}{m}) - 2f'(\delta_{l_1 l_2}) \\
&\quad - f'(\delta_{l_1 l_2} - \frac{2\pi}{m})
\end{aligned} \tag{A.2}$$

where $\delta_{l_1 l_2} \in [0, \frac{2\pi}{m}]$ is the phase shift between the two constellations. Then, if we can show that for all $\delta \in [0, \frac{2\pi}{m}]$ (A.2) is less than zero then for any values of l_1 and l_2 we will have (A.1) satisfied.

Since f is odd and even around $\frac{\pi}{2}$, f' is even and odd around $\frac{\pi}{2}$ and $g_m'(\delta)$ can be rewritten as

$$\begin{aligned}
g_m'(\delta) &= f'(\delta) \\
&\quad + \sum_{1 \le |j| \le \lfloor \frac{k}{2} \rfloor} \left\{ f'(\delta + \frac{2\pi}{m} j) - f'(\delta - sgn(j)\frac{\pi}{m} + \frac{2\pi}{m} j) \right\} \\
&\quad - \left[f'(\delta + \frac{\pi}{m} k) + f'(\delta - \frac{\pi}{m} k) \right] \mathbf{1}_{[k \text{ odd}]}
\end{aligned}$$

where $\mathbf{1}_{[k \text{ odd}]}$ is the indicator function of the event $[k \text{ odd}]$, the sum is over all the integers j with $1 \le |j| \le \lfloor \frac{k}{2} \rfloor$ and $k = \frac{m-1}{2}$

The last term only appears when k is odd and in fact it is easy to show that

it is always negative as the following calculation shows:

$$
\begin{aligned}
&- f'(\delta + \frac{\pi}{m}k) - f'(\delta - \frac{\pi}{m}k) = \\
&= -f'(\frac{\pi}{m}k + \delta) - f'(\frac{\pi}{m}k - \delta) \\
&= -f'(\frac{\pi}{2} - \frac{\pi}{2m} + \delta) - f'(\frac{\pi}{2} - \frac{\pi}{2m} - \delta) \\
&= f'(\frac{\pi}{2} - \delta + \frac{\pi}{2m}) - f'(\frac{\pi}{2} - \delta - \frac{\pi}{2m}) \\
&= f'(\theta) - f'(\theta - \phi) < 0
\end{aligned}
$$

where in step one we used the fact of f' being even, in step two we used $k = \frac{m-1}{2}$ and in step three we use f' being odd around $\frac{\pi}{2}$. The last step comes from substituting $\theta = \frac{\pi}{2} - \delta + \frac{\pi}{2m}$, $\phi = \frac{\pi}{m}$ and apply Lemma 2.3, since for $m \geq 7$ we have $0 \leq \theta - \phi < \theta \leq \pi$.

Then it remains the show that the terms of the form $f'(\delta + \frac{2\pi}{m}j) - f'(\delta - sgn(j)\frac{\pi}{m} + \frac{2\pi}{m}j)$ are negative for all j s.t. $1 \leq |j| \leq \lfloor \frac{k}{2} \rfloor$. This is indeed true when j is positive since for all $\delta \in [0, \frac{2\pi}{m}]$ we get

$$0 \leq \delta - \frac{\pi}{m} + \frac{2\pi}{m}j < \delta + \frac{2\pi}{m}j \leq \pi, \text{ for } 1 \leq j \leq \lfloor \frac{k}{2} \rfloor$$

and thus we can apply again Lemma 2.3.

When j is negative there is one exception in which Lemma 2.3 cannot be used since

$$-\pi \leq \delta + \frac{2\pi}{m}j < \delta + \frac{2\pi}{m}j + \frac{\pi}{m} \leq 0, \forall \delta \in [0, \frac{2\pi}{m}]$$

only holds for $-\lfloor \frac{k}{2} \rfloor \leq j \leq -2$. Thus the term corresponding to $j = -1$ cannot be directly eliminated.

Then, by keeping only the terms of the sum with $j = \pm 1$, g'_m is strictly upper

bounded for all $\delta \in [0, \frac{2\pi}{m}]$ by

$$\begin{aligned} g'_m(\delta) <& f'(\delta) + f'(\delta - \frac{2\pi}{m}) - f'(\delta - \frac{\pi}{m}) \\ &+ f'(\delta + \frac{2\pi}{m}) - f'(\delta + \frac{\pi}{m}) \end{aligned} \tag{A.3}$$

Now substituting (A.3) in (A.2) we get

$$\begin{aligned} &\sum_{ij \in \mathcal{K}_{l_1 l_2}} f'(\phi^*_j - \phi^*_i) \\ &< f'(\delta - \frac{2\pi}{m}) - 2f'(\delta - \frac{\pi}{m}) + f'(\delta + \frac{2\pi}{m}) - 2f'(\delta + \frac{\pi}{m}) \\ &\leq f'(\delta - \frac{2\pi}{m}) - 2f'(\delta - \frac{\pi}{m}) - f'(\delta + \frac{\pi}{m}) \\ &\leq f'(\delta - \frac{2\pi}{m}) - 2f'(\delta - \frac{\pi}{m}) \end{aligned}$$

where in the last step we used the fact that for $m \geq 6$ and $\delta \in [0, \frac{2\pi}{m}]$, $f'(\delta + \frac{\pi}{m}) \geq 0$.

Finally, since for $\delta \in [0, \frac{2\pi}{m}]$ $f'(\delta - \frac{2\pi}{m})$ is strictly increasing and $f'(\delta - \frac{\pi}{m})$ achieves its minimum for $\delta \in \{0, \frac{2\pi}{m}\}$, then

$$f'(\delta - \frac{2\pi}{m}) - 2f'(\delta - \frac{\pi}{m}) \leq f'(0) - 2f'(\frac{\pi}{m}) \leq 0$$

where the last inequality follow from Lemma 2.4.

Therefore, for all m odd greater or equal to 7 we obtain

$$\sum_{ij \in \mathcal{K}_{l_1 l_2}} f'(\phi^*_j - \phi^*_i) < f'(0) - 2f'(\frac{\pi}{m}) \leq 0$$

and since this result is independent on the indices l_1, l_2, then

$$\begin{aligned} &\sum_{ij \in C(S, V(G) \setminus S)} f'(\phi^*_j - \phi^*_i) \\ &\qquad = \sum_{l_1=1}^{l_B} \sum_{l_2=1}^{l_B} \sum_{ij \in \mathcal{K}_{l_1 l_2}} f'(\phi^*_j - \phi^*_i) < 0 \end{aligned}$$

and thus ϕ^* is unstable. □

A.2 Proof of Lemma 3.1

Proof. We first compute the characteristic polynomial

$$\det(\lambda I_{3n} - A) = \begin{vmatrix} (\lambda-1)I_n & -\tau R & 0 \\ \kappa_1 L & (\lambda-1)I_n & \kappa_2 I_n \\ pL & 0 & (\lambda-1+p)I_n \end{vmatrix}$$

$$= (\lambda-1)^n \begin{vmatrix} (\lambda-1)I_n + \frac{\tau\kappa_1}{\lambda-1}LR & \kappa_2 I_n \\ \frac{\tau p}{\lambda-1}LR & (\lambda-1+p)I_n \end{vmatrix}$$

$$= \det\Big((\lambda-1)^2(\lambda-1+p)I_n + [(\lambda-1)\kappa_1$$

$$+(\kappa_2-\kappa_1)]\tau LR\Big) = \prod_{l=1}^{n} g_l(\lambda),$$

where $g_l(\lambda)$ is as defined in (3.20) and we have iteratively use the determinant property of block matrices $\det(A) = \det(A_{11})\det(A\backslash A_{11})$ where $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $A\backslash A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is the Schur complement of A_{11} [150].

Thus, $\lambda = 1$ is a double root of the characteristic polynomial if and only if $\kappa_1 \neq \kappa_2$, $p > 0$ and τcLR has a simple zero eigenvalue, i.e. (3.21). Now, since R is nonsingular (3.21) must hold for the eigenvalues of L as well, which is in fact true if and only if the directed graph $G(V, E)$ is connected [148]. □

A.3 Proof of Lemma 3.2

Proof. We start by computing the right Jordan chain. By definition of ζ_1, $(A - I)\zeta_1 = 0_n$. Thus, if $\zeta_1 = [x^T \ s^T \ y^T]^T$, then the following system of equations must

be satisfied

$$\tau Rs = 0 \text{ (a)}, \ -\kappa_1 Lx - \kappa_2 y = 0 \text{ (b)}, \ -pLx - py = 0 \text{ (c)}. \tag{A.4}$$

Equation (A.4a) implies $s = 0$. Now, since $p > 0$, (A.4c) implies $Lx = -y$, which when substituted in (A.4b) gives $(\kappa_2 - \kappa_1)y = 0$. Thus, since $\kappa_1 \neq \kappa_2$, $y = 0$ and $x \in \ker(L)$. By choosing $x = \alpha_1 \mathbf{1}_n$ (for some $\alpha_1 \neq 0$) we obtain $\zeta_1 = \alpha_1 \left[\mathbf{1}_n^T \ \mathbf{0}_n^T \ \mathbf{0}_n^T\right]^T$.

Notice that the computation also shows that ζ_1 is the unique eigenvector of $\mu(A) = 1$ which implies that there is only one Jordan block of size 2. The second member of the chain, ζ_2, can be computed similarly by solving $(A - I_n)\zeta_2 = \zeta_1$ and $(A - (1-p)I_n)\zeta_3 = 0$ we get

$$\zeta_2 = \begin{bmatrix} \alpha_2 \mathbf{1}_n \\ \frac{\alpha_1}{\tau} R^{-1}\mathbf{1}_n \\ \mathbf{0}_n \end{bmatrix} \quad \text{and} \quad \zeta_3 = \alpha_3 \begin{bmatrix} -\frac{\tau\kappa_2}{p^2}\mathbf{1}_n \\ \frac{\kappa_2}{p} R^{-1}\mathbf{1}_n \\ R^{-1}\mathbf{1}_n \end{bmatrix}.$$

To compute ζ_3, first notice that $(A - (1-p)I_n)\zeta_3 = 0$ implies $Lx =$ and $x = -\frac{\tau}{p}s = -\frac{\kappa_2\tau}{p^2}y$. ζ_3 follows by taking $y = \alpha_3 R^{-1}\mathbf{1}_n$.

The vectors η_1, η_2 and η_3 can be solved in the same way using $\eta_2^T(A - I) = 0$, $\eta_1^T(A - I) = \eta_2^T$ and $\eta_3^T(A - (1-p)I) = 0$. This gives $\eta_1 = \left[\frac{\beta_2}{\tau} R^{-1}\xi^T \ \beta_1\xi^T \ (-\frac{\kappa_2}{p}\beta_1 + \frac{\kappa_2}{p^2}\beta_2)\xi^T\right]^T$, $\eta_2 = \beta_2\left[\mathbf{0}_n^T \ \xi^T \ \frac{\kappa_2}{p}\xi^T\right]^T$ and $\eta_3 = \beta_3\left[\mathbf{0}_n^T \ \mathbf{0}_n^T \ \xi^T\right]^T$. We set $\alpha_1 = \alpha_2 = \alpha_3 = 1$; this can be done WLOG provided we still satisfy $\eta_l^T\zeta_l = 1$ and $\eta_l^T\zeta_h = 0$ for $l \neq h$. Finally, $\eta_1^T\zeta_1 = 1$ gives$\beta_2 = \gamma\tau$, $\eta_3^T\zeta_3 = 1$ gives $\beta_3 = \gamma$ and $\eta_1^T\zeta_2 = 0$ gives $\beta_1 = -\beta_2 = -\gamma\tau$. □

A.4 Proof of Theorem 3.1

Proof. We first notice that whenever $x(t_k)$ approaches (3.17) then

$$\lim_{h\to\infty} x(t_h) - r^*\mathbf{1}_n t_h = x^*\mathbf{1}_n \tag{A.5}$$

Sufficiency

Since we are under the assumptions of Lemmas 3.1 and 3.2 we know that $\mu(A) = 1$ has multiplicity 2 and a Jordan chain of size 2. Thus, the Jordan normal form of A is

$$A = [\zeta_1 \ldots \zeta_{3n}] \begin{bmatrix} 1 & 1 & 0 & & \\ 0 & 1 & 0 & \mathbf{0}_{3\times 3(n-1)} & \\ 0 & 0 & 1-p & & \\ & \mathbf{0}_{3(n-1)\times 3} & & \hat{J} & \end{bmatrix} \begin{bmatrix} {\eta_1}^T \\ \vdots \\ {\eta_{3n}}^T \end{bmatrix} \tag{A.6}$$

where $\hat{J}$ has eigenvalues with spectral radius $\rho(\hat{J}) := \max_l |\mu_l(\hat{J})| < 1$. Thus, it follows that

$$\lim_{h\to\infty} A^h - \zeta_1\eta_1^T - (h\zeta_1 + \zeta_2)\eta_2^T = \lim_{h\to\infty} [\zeta_1 \ldots \zeta_{3n}] \tag{A.7}$$
$$\begin{bmatrix} \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 1} & & \\ \mathbf{0}_{1\times 2} & (1-p)^h & \mathbf{0}_{2\times(3n-2)} & \\ & & & \\ & \mathbf{0}_{(3n-2)\times 2} & \hat{J}^h & \end{bmatrix} \begin{bmatrix} {\eta_1}^T \\ \vdots \\ {\eta_{3n}}^T \end{bmatrix} = \mathbf{0}_{3n}$$

where the last equality follows since $(1-p)^h \xrightarrow[h\to\infty]{} 0$ and $\left\|\hat{J}^h\right\|_\varepsilon \leq \left\|\hat{J}\right\|_\varepsilon^h \leq (\rho + \varepsilon)^h \xrightarrow[h\to\infty]{} 0$, where the norm $\|\cdot\|_\varepsilon$ is chosen such that $\|A\|_\varepsilon = \rho(A) + \varepsilon$ [150, p. 297, Lemma 5.6.10] and ε is such $\rho(\hat{J}) + \varepsilon < 1$.

Right multiplying (A.7) with a given initial condition $z_0 = [x_0^T \; s_0^T \; y_0^T]^T$ and

using (3.22) and (3.23) gives

$$\lim_{k\to\infty} x_k - t_k\gamma\mathbf{1}_n\xi^T(s_0 - \frac{\kappa_2}{p}y_0) = \gamma\mathbf{1}_n\xi^T(R^{-1}x_0 + \tau\frac{\kappa_2}{p^2}y_0). \tag{A.8}$$

Thus, equation (3.24) follows from identifying (A.8) and (A.5).

Necessity

The algorithm achieves synchronization whenever (A.5) holds. Then, it follows from (3.18) and (A.5) that asymptotically the system behaves according to

$$\begin{aligned} z_k = \begin{bmatrix} x_k \\ s_k \\ y_k \end{bmatrix} &= \begin{bmatrix} x^*\mathbf{1}_n \\ r^*R^{-1}\mathbf{1}_n \\ \mathbf{0}_n \end{bmatrix} + k\begin{bmatrix} \tau r^*\mathbf{1}_n \\ \mathbf{0}_n \\ \mathbf{0}_n \end{bmatrix} \\ &= (\tau r^*\zeta_2 + (x^* - \tau r^*)\zeta_1) + kr^*\tau\zeta_2. \end{aligned}$$

Thus, since P is invertible ζ_l are linearly independent. Therefore, if the system synchronizes for arbitrary initial condition, then it must be the case that the effect of the remaining modes $\mu_l(\Gamma)$ vanishes, which can only happen if for every $\mu_l(\Gamma) \neq 1$, $|\mu_l(\Gamma)| < 1$ and the multiplicity of $\mu_l(\Gamma) = 1$ is two. Now suppose that either $\kappa_1 = \kappa_2$ or $p = 0$. Then by Lemma 3.1, the multiplicity of $\mu_l(\Gamma) = 1$ is not two which is a contradiction. Thus, we must have $\kappa_1 \neq \kappa_2$ and $p > 0$ whenever the system synchronizes for arbitrary initial condition. □

A.5 Proof of Theorem 3.3

Proof. We will show that when $G(V,E)$ is connected with $\mu(L) \in \mathbb{R}$, then (i)-(iii) are equivalent to the conditions of Theorem 3.1.

Since, $G(V,E)$ is connected and (i)-(ii) satisfies $p > 0$ and $\kappa_1 \neq \kappa_2$, the conditions of Lemma 3.1 are satisfied. Therefore the multiplicity of $\mu(A) = 1$ is two and by (3.21) these are the roots of $g_n(\lambda) = (\lambda-1)^2(\lambda-1+p)$, which corresponds to the case $\nu_n = 0$.

Thus, to satisfy Theorem 3.1 we need to show that the remaining eigenvalues are strictly in the unit circle. This is true for the remaining root of $g_n(\lambda)$ iff (i).

For the remaining $g_l(\lambda)$, this implies that are Schur polynomials. Thus, we will show that $g_l(\lambda)$ is a Schur polynomial if and only if (i)-(iii) hold. We drop the subindex l for the rest of the proof.

We first transform the Schur stability problem into a Hurwitz stability problem. Consider the change of variable $\lambda = \frac{s+1}{s-1}$. Then $|\lambda| < 1$ if and only if $\mathbb{R}[s] < 0$.

Now, since $\nu > 0$ by (3.21), let

$$P(s) = \frac{(s-1)^3}{\delta\kappa p\nu} g\left(\frac{s+1}{s-1}\right) = s^3 + \left(\frac{2\kappa_1}{\delta\kappa p} - 3\right)s^2 + \left(\frac{4}{\delta\kappa\nu} + 3 - \frac{4\kappa_1}{\delta\kappa p}\right)s + \frac{4(2-p)}{\delta\kappa p\nu} + \frac{2\kappa_1}{\delta\kappa p} - 1$$

where $\delta\kappa = \kappa_1 - \kappa_2$.

We will apply Hermite-Beihler Theorem to $P(s)$, but first let us express what 1) and 2) of Theorem 3.2 mean here.

Condition 1) becomes

$$\frac{2\kappa_1}{\delta\kappa p} - 3 > 0. \tag{A.9}$$

Now let $P^r(\omega)$ and $P^i(\omega)$ be as in Theorem 3.2, i.e. let

$$P^r(\omega) = -\left(\frac{2\kappa_1}{\delta\kappa p} - 3\right)\omega^2 + \frac{4(2-p)}{\delta\kappa p\nu} + \frac{2\kappa_1}{\delta\kappa p} - 1$$
$$P^i(\omega) = -\omega^3 + \left(\frac{4}{\delta\kappa\nu} + 3 - \frac{4\kappa_1}{\delta\kappa p}\right)\omega$$

The roots of $P^r(\omega)$ and $P^i(\omega)$ are given by $\omega_0 = \pm\sqrt{\omega_0^r}$ and $\omega_0 \in \{0, \pm\sqrt{\omega_0^i}\}$ respectively, where

$$\omega_0^r := \frac{\frac{4(2-p)}{\delta\kappa p\nu} + \frac{2\kappa_1}{\delta\kappa p} - 1}{\frac{2\kappa_1}{\delta\kappa p} - 3} \text{ and } \omega_0^i := \frac{4}{\delta\kappa\nu} + 3 - \frac{4\kappa_1}{\delta\kappa p} \tag{A.10}$$

Since the roots $P^r(\omega)$ and $P^i(\omega)$ must be real, we must have $\omega_0^r > 0$ and $\omega_0^i > 0$. Therefore, by monotonicity of the square root, the interlacing condition 2) is equivalent to

$$0 < \omega_0^r < \omega_0^i. \tag{A.11}$$

Thus we will show: (i)-(iii) hold $\iff$ (A.9) and (A.11) hold.

It is straightforward to see that using (i) and (ii) we can get (A.9). On the other hand, $\omega_o^i > 0$ from (A.11) together with (A.9) gives $0 < \frac{4}{\delta\kappa\nu} + 3 - \frac{4\kappa_1}{\delta\kappa p} < \frac{4}{\delta\kappa\nu}$, which implies that $\delta\kappa > 0$, and therefore (ii) follows.

Now using (A.9) and the definition of ω_0^r in (A.10), $\omega_0^r > 0$ becomes $\frac{4(2-p)}{\delta\kappa p\nu} + \frac{2\kappa_1}{\delta\kappa p} - 1 > 0$ which always holds under (i) and (ii) since the first term is always positive and $\frac{2\kappa_1}{\delta\kappa p} - 1 > \frac{2\kappa_1}{\delta\kappa p} - 3 > 0$ by (A.9).

Using (A.10), $\omega_0^r < \omega_0^i$ is equivalent to

$$\frac{\frac{4(2-p)}{\delta\kappa p\nu} + \frac{2\kappa_1}{\delta\kappa p} - 1}{\frac{2\kappa_1}{\delta\kappa p} - 3} < \frac{4}{\delta\kappa\nu} + 3 - \frac{4\kappa_1}{\delta\kappa p}$$
$$\frac{\frac{2\kappa_1}{\delta\kappa p} - 1}{\frac{2\kappa_1}{\delta\kappa p} - 3} + \frac{4\kappa_1}{\delta\kappa p} - 3 < \frac{4}{\delta\kappa\nu}\left(1 - \frac{\frac{(2-p)}{p}}{\frac{2\kappa_1}{\delta\kappa p} - 3}\right)$$

where the left-hand side (LHS) is

$$LHS = \frac{(2\kappa_1 - \delta\kappa p)\delta)\delta\kappa p + (4\kappa_1 - 3\delta\kappa p)(2\kappa_1 - 3\delta\kappa p)}{(2\kappa 1 - 3\delta\kappa p)\delta\kappa p}$$
$$= \frac{8(\kappa_1^2 - 2\kappa_1\delta\kappa p + (\delta\kappa p)^2)}{(2\kappa 1 - 3\delta\kappa p)\delta\kappa p} = \frac{8(\kappa_1 - \delta\kappa p)^2}{(2\kappa 1 - 3\delta\kappa p)\delta\kappa p}$$

and the right hand side (RHS) is

$$RHS = \frac{4}{\delta\kappa\nu}\frac{\frac{2\kappa_1 - 3\delta\kappa p + (2-p)\delta\kappa}{\delta\kappa p}}{\frac{2\kappa_1 - 3\delta\kappa p}{\delta\kappa p}} = \frac{8}{\delta\kappa\nu}\frac{\kappa_2 - \delta\kappa p}{2\kappa_1 - 3\delta\kappa p}.$$

Thus $LHS < RHS$ becomes

$$\begin{aligned} \frac{8(\kappa_1 - \delta\kappa p)^2}{(2\kappa 1 - 3\delta\kappa p)\delta\kappa p} &< \frac{8}{\delta\kappa\nu}\frac{\kappa_2 - \delta\kappa p}{2\kappa_1 - 3\delta\kappa p} \\ \frac{(\kappa_1 - \delta\kappa p)^2}{p} &< \frac{1}{\nu}(\kappa_2 - \delta\kappa p) \\ \nu &< \frac{p(\kappa_2 - \delta\kappa p)}{(\kappa_1 - \delta\kappa p)^2} \end{aligned} \tag{A.12}$$

Finally, $\nu_l = \mu_l(\tau LR) = \tau\mu_l(LR)$. Thus, since (A.12) should hold $\forall l \in \{1, ..., n-1\}$, then

$$\tau < \min_l \frac{p(\kappa_2 - \delta\kappa p)}{\mu_l(LR)(\kappa_1 - \delta\kappa p)^2} = \frac{p(\kappa_2 - \delta\kappa p)}{\mu_{\max}(\kappa_1 - \delta\kappa p)^2}$$

which is exactly (iii). □

A.6 Graph Laplacian with Real Eigenvalues

We know show that every graph G with a leader i_0 such that the graph G_{sub} induced by $V\backslash\{1\}$ is symmetric, always has real eigenvalues. WLOG assume $i_0 = 1$. Then from the structure of the graph it is easy to see that

$$L = \left[\begin{array}{c|c} 0 & \mathbf{0}_{n-1}^T \\ \hline l & \tilde{L} \end{array}\right]$$

where $l_{i-1} = -\alpha_{i1}$,

$$\tilde{L} = L_{G_{\text{sub}}} + D_{n-1}$$

and $D_{n-1} = \text{diag}[\alpha_{i1}]$. Thus, since $L_{G_{\text{sub}}}$ is a symmetric matrix and D_{n-1} diagonal, it follows that the eigenvalues of $\tilde{L}$ are real. Finally consider a possibly complex eigenvalue λ and corresponding eigenvector $x = [x^1|(x^{[2,n]})^T]^T$. Then, since $Lx = \lambda x$, it follows that

$$0 = \mathbf{0}_n^T x = \lambda x^1 \quad \text{and} \quad l x^1 + \tilde{L} x^{[2,n]} = \lambda x^{[2,n]}.$$

Thus, whenever $\lambda \neq 0$, $x^1 = 0$ and thus we obtain $\tilde{L} x^{[2,n]} = \lambda x^{[2,n]}$ which implies that λ is an eigenvalue of $\tilde{L}$. This show our claim since we have already proved that $\tilde{L}$ as symmetric and therefore can only have real eigenvalues. □

BIBLIOGRAPHY

[1] A. T. Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, 16(1):15 – 42, 1967.

[2] C. S. Peskin. *Mathematical aspects of heart physiology*. Courant Institute of Mathematical Sciences, New York University, New York, NY, USA, 1975.

[3] Peter Achermann and Hanspeter Kunz. Modeling circadian rhythm generation in the suprachiasmatic nucleus with locally coupled self-sustained oscillators: Phase shifts and phase response curves. *Journal of Biological Rhythms*, 14(6):460 – 468, 1999.

[4] Jordi Garcia-Ojalvo, Michael B. Elowitz, Steven H. Strogatz, and Charles S. Peskin. Modeling a synthetic multicellular clock: Repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30):10955–10960, 2004.

[5] S. Yamaguchi, H. Isejima, T. Matsuo, R. Okura, K. Yagita, M. Kobayashi, and H. Okamura. Synchronization of cellular clocks in the suprachiasmatic nucleus. *Science*, 32:1408–1412, Nov. 2003.

[6] I. Z. Kiss, Y. Zhai, and J. L. Hudson. Emerging coherence in a population of chemical oscillators. *Science*, 296:1676–1678, Nov. 2002.

[7] R.A. York and R.C. Compton. Quasi-optical power combining using mutually synchronized oscillator arrays. *IEEE Transactions on Microwave Theory and Techniques*, 39(6):1000–1009, Jun. 1991.

[8] S. A. Marvel and S. H. Strogatz. Invariant submanifold for series arrays of josephson junctions. *Chaos*, 19(1):013132, Mar. 2009.

[9] P. C. Bressloff and S. Coombes. Travelling waves in chains of pulse-coupled integrate-and-fire oscillators with distributed delays. *Phys. D*, 130(3-4):232–254, 1999.

[10] G. B. Ermentrout. Stable periodic solutions to discrete and continuum arrays of weakly coupled nonlinear oscillators. *SIAM J. Appl. Math.*, 52(6):1665–1687, 1992.

[11] Roger D Traub and RK Wong. Cellular mechanism of neuronal synchronization in epilepsy. *Science*, 216(4547):745–747, 1982.

[12] David A McCormick and Diego Contreras. On the cellular and network bases of epileptic seizures. *Annual Review of Physiology*, 63(1):815–846, 2001.

[13] Theoden I Netoff and Steven J Schiff. Decreased neuronal synchronization during experimental seizures. *The Journal of neuroscience*, 22(16):7297–7307, 2002.

[14] Yuji Nakayama and Satoru Aikawa. Cell discard and tdma synchronization using fec in wireless atm systems. *Selected Areas in Communications, IEEE Journal on*, 15(1):29–34, 1997.

[15] Vilgot Claesson, Henrik Lonn, and Neeraj Suri. Efficient tdma synchronization for distributed embedded systems. In *Reliable Distributed Systems, 2001. Proceedings. 20th IEEE Symposium on*, pages 198–201. IEEE, 2001.

[16] W. Ye, J. Heidemann, and D. Estrin. Medium access control with coordinated adaptive sleeping for wireless sensor networks. *IEEE/ACM Trans. Netw.*, 12:493–506, June 2004.

[17] Yao-Win Hong and A. Scaglione. A scalable synchronization protocol for large scale sensor networks and its applications. *IEEE Journal on Selected Areas in Communications*, 23(5):1085–1099, May. 2005.

[18] Geoffrey Werner-Allen, Geetika Tewari, Ankit Patel, Matt Welsh, and Radhika Nagpal. Firefly-inspired sensor network synchronicity with realistic radio effects. In *SenSys: Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, pages 142–153, New York, NY, USA, 2005. ACM.

[19] Dan Li, K.D. Wong, Yu Hen Hu, and A.M. Sayeed. Detection, classification, and tracking of targets. *Signal Processing Magazine, IEEE*, 19(2):17 –29, March 2002.

[20] J.N. Laneman and G.W. Wornell. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. *Information Theory, IEEE Transactions on*, 49(10):2415 – 2425, oct. 2003.

[21] N Wiener. Nonlinear problems in random theory, 1958.

[22] Arthur T Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, 16(1):15–42, 1967.

[23] Y. Kuramoto. International symposium on mathematical problems in theoretical physics. In *Lecture notes in Physics*, volume 39, page 420. Springer, 1975.

[24] R. E. Mirollo and S. H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math*, 50:1645–1662, 1990.

[25] E. M. Izhikevich. Weakly pulse-coupled oscillators, fm interactions, synchronization, and oscillatory associative memory. *IEEE Transactions on Neural Networks*, 10(3):508–526, May. 1999.

[26] LL Bonilla, CJ Perez Vicente, and R Spigler. Time-periodic phases in populations of nonlinearly coupled oscillators with bimodal frequency distributions. *Physica D: Nonlinear Phenomena*, 113(1):79–97, 1998.

[27] Steven H Strogatz and Renato E Mirollo. Stability of incoherence in a population of coupled oscillators. *Journal of Statistical Physics*, 63(3-4):613–635, 1991.

[28] Edward Ott and Thomas M Antonsen. Long time evolution of phase oscillator systems. *Chaos: An interdisciplinary journal of nonlinear science*, 19(2):023117–023117, 2009.

[29] EA Martens, E Barreto, SH Strogatz, E Ott, P So, and TM Antonsen. Exact results for the kuramoto model with a bimodal frequency distribution. *Physical Review E*, 79(2):026204, 2009.

[30] Y. Kuramoto. Cooperative dynamics of oscillator community. In *Progress of Theoretical Physics Supplement*, volume 79, 1984.

[31] Steven H Strogatz. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20, 2000.

[32] P. Monzón and F. Paganini. Global considerations on the kuramoto model of sinusoidally coupled oscillators. In *Proceedings of the 44th IEEE Conference on Decision and Control, and European Control Conference*, pages 3923–3928, Sevilla, Spain, Dec. 2005.

[33] P. C. Bressloff and S. Coombes. Symmetry and phase-locking in a ring of pulse-coupled oscillators with distributed delays. *Phys. D*, 126(1-2):99–122, 1999.

[34] Eduardo Canale and Pablo Monzón. Gluing kuramoto coupled oscillators networks. In *Decision and Control, 2007 46th IEEE Conference on*, pages 4596–4601. IEEE, 2007.

[35] Eduardo A Canale and Pablo A Monzon. On the characterization of families of synchronizing graphs for kuramoto coupled oscillators. In *Estimation and Control of Networked Systems*, volume 1, pages 42–47, 2009.

[36] P. Ashwin and J. W. Swift. The dynamics of n weakly coupled identical oscillators. *J. Nonlinear Sci*, 2(1):69–108, 1992.

[37] O. V. Popovych, Y. L. Maistrenko, and P. A. Tass. Phase chaos in coupled oscillators. *Phys. Rev. E*, 71(6):065201, Jun. 2005.

[38] E. Brown, P. Holmes, and J. Moehlis. Globally coupled oscillator networks. In *Perspectives and Problems in Nonlinear Science: A Celebratory Volume in Honor of Larry Sirovich*, pages 183–215. Springer, 2003.

[39] A. Jadbabaie, N. Motee, and M. Barahona. On the stability of the kuramoto model of coupled nonlinear oscillators. In *Proceedings of the American Control Conference.*, volume 5, pages 4296–4301, June 30, 2004.

[40] D. Lucarelli and I-Jeng Wang. Decentralized synchronization protocols with nearest neighbor communication. In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, 2004.

[41] David Mills. Network time protocol version 4 reference and implementation guide. Technical Report 06-06-1, University of Delaware, June 2006.

[42] Ahmed Sobeih, Michel Hack, Zhen Liu, and null Li Zhang. Almost peer-to-peer clock synchronization. *Parallel and Distributed Processing Symposium, International*, page 21, 2007.

[43] IEEE standard for a precision clock synchronization protocol for networked measurement and control systems. pages 1 –269, 2008.

[44] Darryl Veitch, Julien Ridoux, and Satish Babu Korada. Robust synchronization of absolute and difference clocks over networks. *IEEE/ACM Trans. Netw.*, 17(2):417–430, April 2009.

[45] R. Carli and S. Zampieri. Networked clock synchronization based on sec-

ond order linear consensus algorithms. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 7259 –7264, December 2010.

[46] E. Mallada and Ao Tang. Distributed clock synchronization: Joint frequency and phase consensus. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 6742 –6747, December 2011.

[47] J. Ridoux, D. Veitch, and T. Broomhead. The case for feed-forward clock synchronization. *IEEE/ACM Transactions on Networking*, 20(1):231–242, 2012.

[48] James C. Corbett, Jeffrey Dean, Michael Epstein, Andrew Fikes, Christopher Frost, J. J. Furman, Sanjay Ghemawat, Andrey Gubarev, Christopher Heiser, Peter Hochschild, Wilson Hsieh, Sebastian Kanthak, Eugene Kogan, Hongyi Li, Alexander Lloyd, Sergey Melnik, David Mwaura, David Nagle, Sean Quinlan, Rajesh Rao, Lindsay Rolig, Yasushi Saito, Michal Szymaniak, Christopher Taylor, Ruth Wang, and Dale Woodford. Spanner: Google's globally-distributed database. In *Proceedings of the 10th USENIX conference on Operating Systems Design and Implementation*, OSDI'12, pages 251–264, Berkeley, CA, USA, 2012. USENIX Association.

[49] S. Froehlich, M. Hack, Xiaoqiao Meng, and Li Zhang. Achieving precise coordinated cluster time in a cluster environment. In *Precision Clock Synchronization for Measurement, Control and Communication, 2008. ISPCS 2008. IEEE International Symposium on*, pages 54 –58, September 2008.

[50] Li Zhang, Zhen Liu, and C. Honghui Xia. Clock synchronization algorithms for network measurements. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 160 – 169 vol.1, 2002.

[51] Hayang Kim, Xiaoli Ma, and B.R. Hamilton. Tracking low-precision clocks with time-varying drifts using kalman filtering. *Networking, IEEE/ACM Transactions on*, 20(1):257 –270, February 2012.

[52] Jeremy Elson, Lewis Girod, and Deborah Estrin. Fine-grained network time synchronization using reference broadcasts. *SIGOPS Oper. Syst. Rev.*, 36(SI):147–163, 2002.

[53] D. Hunt, G Korniss, and B Szymanski. Network synchronization in a noisy environment with time delays: Fundamental limits and trade-offs. *Physical Review Letters*, 105(6):068701, 2010.

[54] O. Gurewitz, I. Cidon, and M. Sidi. Network classless time protocol based on clock offset optimization. *IEEE/ACM Transactions on Networking*, 14(4):876–888, 2006.

[55] J.-M. Berthaud. Time synchronization over networks using convex closures. *IEEE/ACM Transactions on Networking*, 8(2):265–277, 2000.

[56] J. Escobar, C. Partridge, and D. Deutsch. Flow synchronization protocol. *IEEE/ACM Transactions on Networking*, 2(2):111–121, 1994.

[57] D. L. Mills. Internet time synchronization: the network time protocol. *IEEE Transactions on Communications*, 39(10):1482–1493, 1991.

[58] Hicham Marouani and Michel R. Dagenais. Internal clock drift estimation in computer clusters. *J. Comp. Sys., Netw., and Comm.*, 2008:9:1–9:7, January 2008.

[59] S.B. Moon, P. Skelly, and D. Towsley. Estimation and removal of clock skew from network delay measurements. In *INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 227 –234 vol.1, March 1999.

[60] M.D. Lemmon, J. Ganguly, and L. Xia. Model-based clock synchronization in networks with drifting clocks. In *Dependable Computing, 2000. Proceedings. 2000 Pacific Rim International Symposium on*, pages 177 –184, 2000.

[61] Massoud Amin. Epri/dod complex interactive networks/systems initiative: Self-healing infrastructures. *B PJOSBBL*, page 3, 2000.

[62] R. Sugarman. *New York City's blackout: a $350 million drain*. 1978.

[63] Douglas W. Caves, Joseph A. Herriges, and Robert J. Windle. The cost of electric power interruptions in the industrial sector: Estimates derived from interruptible service programs. *Land Economics*, 68(1):49–61, February 1992. ArticleType: research-article / Full publication date: Feb., 1992 / Copyright © 1992 The Board of Regents of the University of Wisconsin System.

[64] A. F. Adenikinju. *Analysis of the cost of infrastructure failures in a developing economy: The case of the electricity sector in Nigeria*, volume 148. African Economic Research Consortium, 2005.

[65] S. K. Joo, J. C. Kim, and C. C. Liu. Empirical analysis of the impact of 2003 blackout on security values of us utilities and electrical equipment manufacturing firms. *Power Systems, IEEE Transactions on*, 22(3):1012–1018, 2007.

[66] C. A. Cañizares. Applications of optimization to voltage collapse analysis. In *IEEE-PES Summer Meeting, San Diego, USA*, 1998.

[67] I. Dobson. Distance to bifurcation in multidimensional parameter space: Margin sensitivity and closest bifurcations. *Bifurcation Control*, pages 704–706, 2003.

[68] S. Greene, I. Dobson, and F. L. Alvarado. Sensitivity of the loading margin to voltage collapse with respect to arbitrary parameters. *Power Systems, IEEE Transactions on*, 12(1):262–272, 1997.

[69] M. Klein, G.J. Rogers, and P. Kundur. A fundamental study of inter-area oscillations in power systems. *IEEE Transactions on Power Systems*, 6(3):914 –921, August 1991.

[70] M. Klein, L.X. Le, G.J. Rogers, S. Farrokhpay, and N.J. Balu. H infin; damping controller design in large power systems. *IEEE Transactions on Power Systems*, 10(1):158 –166, February 1995.

[71] R.A. Jabr, B.C. Pal, N. Martins, and J.C.R. Ferraz. Robust and coordinated tuning of power system stabiliser gains using sequential linear programming. *IET Generation, Transmission Distribution*, 4(8):893 –904, August 2010.

[72] A. H. El-Abiad and K. Nagappan. Transient stability regions of multimachine power systems. *Power Apparatus and Systems, IEEE Transactions on*, (2):169–179, 1966.

[73] T. Athay, R. Podmore, and S. Virmani. A practical method for the direct analysis of transient stability. *Power Apparatus and Systems, IEEE Transactions on*, (2):573–584, 1979.

[74] A. A. Fouad and V. Vittal. *Power system transient stability analysis using the transient energy function method*. Prentice Hall, 1991.

[75] GC Ejebe, GD Irisarri, S Mokhtari, O Obadina, P Ristanovic, and J Tong. Methods for contingency screening and ranking for voltage stability anal-

ysis of power systems. In *Power Industry Computer Application Conference, 1995. Conference Proceedings., 1995 IEEE*, pages 249–255. IEEE, 1995.

[76] Hsiao-Dong Chiang, Cheng-Shan Wang, and Alexander J Flueck. Look-ahead voltage and load margin contingency selection functions for large-scale power systems. *Power Systems, IEEE Transactions on*, 12(1):173–180, 1997.

[77] H. D. Chiang, W. Ma, R. J. Thomas, and J. S. Thorp. A tool for analyzing voltage collapse in electric power systems PSCC. 1990.

[78] V. Ajjarapu and C. Christy. The continuation power flow: a tool for steady state voltage stability analysis. *Power Systems, IEEE Transactions on*, 7(1):416–423, 1992.

[79] Hsiao-Dong Chiang, Alexander J Flueck, Kirit S Shah, and Neal Balu. Cpflow: A practical tool for tracing power system steady-state stationary behavior due to load and generation variations. *Power Systems, IEEE Transactions on*, 10(2):623–634, 1995.

[80] E. V. Larsen and D. A. Swann. Applying power system stabilizers part III: practical considerations. *Power Apparatus and Systems, IEEE Transactions on*, (6):3034–3046, 1981.

[81] P. Kundur, M. Klein, G. J. Rogers, and M. S. Zywno. Application of power system stabilizers for enhancement of overall system stability. *Power Systems, IEEE Transactions on*, 4(2):614–626, 1989.

[82] N. Martins and L. T. G. Lima. Determination of suitable locations for power system stabilizers and static var compensators for damping electromechanical oscillations in large scale power systems. *Power Systems, IEEE Transactions on*, 5(4):1455–1469, 1990.

[83] Y.L. Abdel-Magid, M.A. Abido, S. Al-Baiyat, and A.H. Mantawy. Simultaneous stabilization of multimachine power systems via genetic algorithms. *IEEE Transactions on Power Systems*, 14(4):1428 –1439, November 1999.

[84] M. A. Abido. Optimal design of power-system stabilizers using particle swarm optimization. *Energy Conversion, IEEE Transactions on*, 17(3):406–413, 2002.

[85] H. Ni, G. T. Heydt, and L. Mili. Power system stability agents using robust wide area control. *Power Systems, IEEE Transactions on*, 17(4):1123–1131, 2002.

[86] X. Xiaorong, X. Jinyu, T. Luyuan, and H. Yingduo. Inter-area damping control of interconnected power systems using wide-area measurements [j]. *Automation of Electric Power Systems*, 2(28):37–40, 2004.

[87] H. Breulmann, E. Grebe, and M. Lösing. Analysis and damping of inter-area oscillations in the UCTE/CENTREL power system. 2000.

[88] R. Grunbaum and J. Pernot. Thyristor controlled series compensation: A state of the art approach for optimization of transmission over power links. *ABB Review*, (13), 2001.

[89] B. Chaudhuri, B.C. Pal, A.C. Zolotas, I.M. Jaimoukha, and T.C. Green. Mixed-sensitivity approach to h infin; control of power system oscillations employing multiple FACTS devices. *IEEE Transactions on Power Systems*, 18(3):1149 – 1156, August 2003.

[90] C. K. Tang, C. E. Graham, M. El-Kady, and R. T H Alden. Transient stability index from conventional time domain simulation. *IEEE Transactions on Power Systems*, 9(3):1524–1530, August 1994.

[91] H. D. Chiang, F. F. Wu, and P. P. Varaiya. A BCU method for direct analysis of power system transient stability. *Power Systems, IEEE Transactions on*, 9(3):1194–1208, 1994.

[92] Hsiao-Dong Chiang. *Direct Methods for Stability Analysis of Electric Power Systems: Theoretical Foundation, BCU Methodologies, and Applications*. Wiley. com, 2011.

[93] J. Carpentier. Contribution to the economic dispatch problem. *Bulletin Society Francaise Electriciens*, 3(8):431–447, 1962.

[94] R. A. Jabr, A. H. Coonick, and B. J. Cory. A primal-dual interior point method for optimal power flow dispatching. *Power Systems, IEEE Transactions on*, 17(3):654–662, 2002.

[95] Hongye Wang, C.E. Murillo-Sanchez, R.D. Zimmerman, and R.J. Thomas. On computational issues of market-based optimal power flow. *IEEE Transactions on Power Systems*, 22(3):1185 –1193, August 2007.

[96] A. Santos Jr and G. R. M. Da Costa. Optimal-power-flow solution by newton's method applied to an augmented lagrangian function. In *Generation, Transmission and Distribution, IEE Proceedings-*, volume 142, pages 33–36, 1995.

[97] L.S. Vargas, V.H. Quintana, and A. Vannelli. A tutorial description of an interior point method and its applications to security-constrained economic dispatch. *IEEE Transactions on Power Systems*, 8(3):1315 –1324, August 1993.

[98] M. Pavella and P. G. Murthy. Transient stability of power systems: theory and practice. 1994.

[99] A.K. David and Xujun Lin. Dynamic security enhancement in power-market systems. *IEEE Transactions on Power Systems*, 17(2):431 –438, May 2002.

[100] Yong-Hua Song and Xi-Fan Wang. *Operation of Market-oriented Power Systems*. Springer, July 2003.

[101] C.A. Canizares and S.K.M. Kodsi. Power system security in market clearing and dispatch mechanisms. In *IEEE Power Engineering Society General Meeting, 2006*, page 6 pp., 2006.

[102] R. Zarate-Minano, F. Milano, and A.J. Conejo. An OPF methodology to ensure small-signal stability. *IEEE Transactions on Power Systems*, 26(3):1050 –1061, August 2011.

[103] W. Rosehart, C. Cañizares, and V. Quintana. Optimal power flow incorporating voltage collapse constraints. In *Power Engineering Society Summer Meeting, 1999. IEEE*, volume 2, pages 820–825, 1999.

[104] W. Rosehart, C. Cañizares, and V. Quintana. Costs of voltage security in electricity markets. In *Power Engineering Society Summer Meeting, 2000. IEEE*, volume 4, pages 2115–2120, 2000.

[105] C. Canizares, W. Rosehart, A. Berizzi, and C. Bovo. Comparison of voltage security constrained optimal power flow techniques. In *Power Engineering Society Summer Meeting, 2001*, volume 3, pages 1680 –1685 vol.3, 2001.

[106] W. D. Rosehart, C. A. Canizares, and V. H. Quintana. Multiobjective op-

timal power flows to evaluate voltage security costs in power networks. *Power Systems, IEEE Transactions on*, 18(2):578–587, 2003.

[107] A.J. Conejo, F. Milano, and R. Garcia-Bertrand. Congestion management ensuring voltage stability. In *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1 –8, July 2008.

[108] E. Vaahedi, Y. Mansour, C. Fuchs, S. Granville, M.D.L. Latore, and H. Hamadanizadeh. Dynamic security constrained optimal power flow/VAr planning. *IEEE Transactions on Power Systems*, 16(1):38 –43, February 2001.

[109] A.O. Ekwue, H.B. Wan, D.T.Y. Cheng, and Y.H. Song. Singular value decomposition method for voltage stability analysis on the national grid system (NGC). *International Journal of Electrical Power & Energy Systems*, 21(6):425–432, August 1999.

[110] T.B. Nguyen and M.A. Pai. Dynamic security-constrained rescheduling of power systems using trajectory sensitivities. *IEEE Transactions on Power Systems*, 18(2):848–854, May 2003.

[111] C.Y. Chung, Lei Wang, F. Howell, and P. Kundur. Generation rescheduling methods to improve power transfer capability constrained by small-signal stability. *IEEE Transactions on Power Systems*, 19(1):524 – 530, February 2004.

[112] J. Condren and T.W. Gedra. Expected-security-cost optimal power flow with small-signal stability constraints. *IEEE Transactions on Power Systems*, 21(4):1736 –1743, November 2006.

[113] L. Moreau. Stability of continuous-time distributed consensus algorithms. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 4, pages 3998 – 4003 Vol.4, dec. 2004.

[114] Wei Ren and R.W. Beard. Consensus of information under dynamically changing interaction topologies. In *American Control Conference, 2004. Proceedings of the 2004*, volume 6, pages 4939 –4944 vol.6, 30 2004-july 2 2004.

[115] G.-B. Stan and R. Sepulchre. Dissipativity characterization of a class of oscillators and networks of oscillators. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 4, pages 4169 – 4173 vol.4, dec. 2003.

[116] L. Scardovi, A. Sarlette, and R. Sepulchre. Synchronization and balancing on the n-torus. *Systems & Control Letters*, 56(5):335 – 341, 2007.

[117] R. Sepulchre, D. A. Paley, and N. E. Leonard. Stabilization of planar collective motion with limited communication. *Automatic Control, IEEE Transactions on*, 53(3):706–719, 2008.

[118] A. Papachristodoulou and A. Jadbabaie. Synchonization in oscillator networks with heterogeneous delays, switching topologies and nonlinear dynamics. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 4307–4312, Dec. 2006.

[119] A.R. Bergen and D.J. Hill. A structure preserving model for power system stability analysis. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(1):25–35, Jan. 1981.

[120] S. Boyd. Convex optimization of graph laplacian eigenvalues. In *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, pages 1311–1320, 2006.

[121] John M. Lee. *Intorduction to Smooth Manifolds*. Springer, 2003.

[122] L. N. Trefethen. Pseudospectra of linear operators. *Siam Review*, 39(3):383–406, 1997.

[123] J. V. Burke, A. S. Lewis, and M. L. Overton. Two numerical methods for optimizing matrix stability. *Linear Algebra and its Applications*, 351:117–145, 2002.

[124] E. M. Izhikevich. Phase models with explicit time delays. *Phys. Rev. E*, 58(1):905–908, Jul. 1998.

[125] U. Ernst, K. Pawelzik, and T. Geisel. Synchronization induced by temporal delays in pulse-coupled oscillators. *Phys. Rev. Lett.*, 74(9):1570–1573, Feb. 1995.

[126] U. Ernst, K. Pawelzik, and T. Geisel. Delay-induced multistable synchronization of biological oscillators. *Phys. Rev. E*, 57(2):2150–2162, Feb. 1998.

[127] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, Berin Heidelberg New York Tokyo, 1984.

[128] R. Sepulchre, D. A. Paley, and N. E. Leonard. Stabilization of planar collective motion: All-to-all communication. *Automatic Control, IEEE Transactions on*, 52(5):811–824, 2007.

[129] Hiroaki Daido. Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions. *Phys. Rev. Lett.*, 68(7):1073–1076, Feb 1992.

[130] A. Rantzer. A dual to lyapunov's stability theorem. *Systems and Control Letters*, 42:161–168, 2001.

[131] B. Bollobas. *Modern Graph Theory*. Springer, New York, 1998.

[132] C. Godsil and G. Royle. *Algebraic Graph Theory*. Springer, New York, 2001.

[133] H. K. Khalil. *Nonlinear systems; 3rd ed.* Prentice-Hall, 1996.

[134] F. C. Hoppensteadt and E. M. Izhikevich. Weakly connected neural networks, 1997.

[135] Alain Sarlette. *Geometry and symmetries in coordination control*. PhD thesis, University of Liège, Belgium, January 2009.

[136] C. van Vreeswijk, L. Abbott, and G. B. Ermentrout. When inhibition not excitation synchronizes neural firing. *Journal of Computational Neuroscience*, 1(4):313–321, 1994.

[137] Wulfram Gerstner. Rapid phase locking in systems of pulse-coupled oscillators with delays. *Phys. Rev. Lett.*, 76(10):1755–1758, Mar 1996.

[138] S. Watanabe and S. H. Strogatz. Constants of motion for superconducting josephson arrays. *Phys. D*, 74(3-4):197–253, 1994.

[139] Edward Ott and Thomas M. Antonsen. Long time evolution of phase oscillator systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(2):023117, 2009.

[140] Edward Ott and Thomas M. Antonsen. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):037113, 2008.

[141] J.S. Shamma and G. Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *Automatic Control, IEEE Transactions on*, 50(3):312 – 327, march 2005.

[142] E. Mallada and F. Paganini. Stability of node-based multipath routing and dual congestion control. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 1398 –1403, dec. 2008.

[143] D. Goldin, S.A. Attia, and J. Raisch. Consensus for double integrator dynamics with heterogeneous communication topologies. In *Decision and Control, 2010. CDC 2010. 49th IEEE Conference on*, dec. 2010.

[144] Randy A. Freeman. A global attractor consisting of exponentially unstable equilibria. In *Proceedings of American Control Conference*, 2013.

[145] David Carlson and Hans Schneider. Inertia theorems for matrices: The semidefinite case. *Journal of Mathematical Analysis and Applications*, 6(3):430–446, 1963.

[146] E. Mallada and A. Tang. Synchronization of phase-coupled oscillators with arbitrary topology. In *Proceedings of American Control Conference*, 2010.

[147] D. Mills. Network time protocol (version 3) specification, implementation and analysis. 1992.

[148] D. Xie and S. Wang. Consensus of second-order discrete-time multi-agent systems with fixed topology. *Journal of Mathematical Analysis and Applications*, 2011.

[149] S.P. Bhattacharyya, H. Chapellat, and L.H. Keel. *Robust control*. Prentice-Hall Upper Saddle River, New Jersey, 1995.

[150] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge Univ Pr, 1990.

[151] D. Arzelier, D. Georgia, S. Gumussoy, and D. Henrion. H2 for HIFOO. *arXiv preprint arXiv:1010.1442*, 2010.

[152] El-Sayed ME Mostafa. Computational design of optimal discrete-time output feedback controllers. *Journal of the Operations Research Society of Japan*, 51(1):15, 2008.

[153] S. Gumussoy, D. Henrion, M. Millstone, and M. L. OVERTON. Multi-objective robust control with HIFOO 2.0. *arXiv preprint arXiv:0905.3229*, 2009.

[154] Andrey P. Popov, Herbert Werner, and Marc Millstone. Fixed-structure discrete-time controller synthesis with HIFOO. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 3152–3155, 2010.

[155] I. Dobson and L. Lu. New methods for computing a closest saddle node bifurcation and worst case load power margin for voltage collapse. *Power Systems, IEEE Transactions on*, 8(3):905–913, 1993.

[156] C. A. Canizares. Calculating optimal system parameters to maximize the distance to saddle-node bifurcations. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, 45(3):225–237, 1998.

[157] C. A. Canizares, A. Berizzi, and P. Marannino. Using FACTS controllers to maximize available transfer capability. *Proc. Bulk Power Systems Dynamics and Control IV-Restructuring*, pages 633–641, 1998.

[158] Chia-Chi Chu. *Transient dynamics of electric power systems: Direct stability assessment and chaotic motions*. Ph.D., Cornell University, United States – New York, 1996.

[159] Florian Dorfler and Francesco Bullo. Kron reduction of graphs with applications to electrical networks. *arXiv preprint arXiv:1102.2950*, 2011.

[160] F. Dorfler and F. Bullo. Transient stability analysis in power networks and synchronization of non-uniform Kuramoto oscillators. pages 930–937, Baltimore, MD, June 2010.

[161] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, 1982.

[162] I. Dobson. Observations on the geometry of saddle node bifurcation and voltage collapse in electrical power systems. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, 39(3):240 –243, mar. 1992.

[163] Francoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.

[164] A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic theory*. Oxford University Press New York, 1995.

[165] J.-B. Hiriart-Urruty and A.S. Lewis. The clarke and michel-penot subdifferentials of the eigenvalues of a symmetric matrix. *Computational Optimization and Applications*, 13:13–23, 1999. 10.1023/A:1008644520093.

[166] D.J. Gotham and G.T. Heydt. Power flow control and power flow studies for systems with facts devices. *Power Systems, IEEE Transactions on*, 13(1):60 –65, feb. 1998.

[167] Arpita Ghosh, Stephen Boyd, and Amin Saberi. Minimizing effective resistance of a graph. *SIAM Rev.*, 50(1):37–66, 2008.

[168] Kemin Zhou, John Comstock Doyle, and Keith Glover. *Robust and optimal control*, volume 40. Prentice Hall Upper Saddle River, NJ, 1996.

[169] J. V. Burke, A. S. Lewis, and M. L. Overton. Optimization and pseudospectra, with applications to robust stability. *SIAM Journal on Matrix Analysis and Applications*, 25(1):80–104, 2003.

[170] Joe H. Chow and Kwok W. Cheung. A toolbox for power system dynamics and control engineering education and research. *Power Systems, IEEE Transactions on*, 7(4):1559–1564, 1992.

[171] Daniel Kressner and Bart Vandereycken. Subspace methods for computing the pseudospectral abscissa and the stability radius. 2012.

[172] M. A. Pai. *Energy function analysis for power system stability*. Springer, 1989.

[173] Dionysios Barmpoutis and Richard M Murray. Noise propagation in biological and chemical reaction networks. *arXiv preprint arXiv:1108.2538*, 2011.

[174] Dionysios Barmpoutis. *Network structure optimization with applications to minimizing variance and crosstalk*. PhD thesis, California Institute of Technology, 2012.

[175] Ali Jadbabaie, Jie Lin, and A Stephen Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *Automatic Control, IEEE Transactions on*, 48(6):988–1001, 2003.

[176] Reza Olfati-Saber and Richard M Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9):1520–1533, 2004.

[177] Wei Ren and Randal W Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *Automatic Control, IEEE Transactions on*, 50(5):655–661, 2005.

[178] Feng Xiao, Long Wang, and Aiping Wang. Consensus problems in discrete-time multiagent systems with fixed topology. *Journal of mathematical analysis and applications*, 322(2):587–598, 2006.

[179] Xiao Y Wang, Rajeev K Dokania, and Alyssa Apsel. Pco-based synchronization for cognitive duty-cycled impulse radio sensor networks. *Sensors Journal, IEEE*, 11(3):555–564, 2011.